GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2004, 17:06:18 ; Search time 165.707 Seconds
(without alignments)
513.066 Million cell updates/sec

Title: US-10-073-301A-9
Perfect score: 1272
Sequence: 1 QVKLQESGGGLVKPGGSLKL.....TYYCQEWSGYPYTFGGGTKL 237

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1272	100.0	237	7	AAE38657	Aae38657 Mouse G1
2	984	77.4	281	6	AAE33333	Aae33333 12E antib
3	977.5	76.8	255	3	AAB11398	Aab11398 E. coli e
4	977.5	76.8	255	4	AAB74199	Aab74199 PelB-scFv
5	977.5	76.8	255	4	AAB70769	Aab70769 Expressio
6	977.5	76.8	255	4	AAY72020	Aay72020 E. caroto
7	976.5	76.8	316	2	AAY32086	Aay32086 Single ch
8	964.5	75.8	241	6	ABR62010	Abr62010 Single-ch
9	949.5	74.6	241	5	AAM48925	Aam48925 scFv anti
10	947.5	74.5	236	2	AAR32842	Aar32842 VH NQ10/1
11	936	73.6	249	4	AAB20436	Aab20436 Anti-FIX/
12	935.5	73.5	240	2	AAR68613	Aar68613 Single ch
13	932.5	73.3	556	2	AAW90218	Aaw90218 Bispecifi
14	932.5	73.3	580	2	AAW90217	Aaw90217 Bispecifi
15	926.5	72.8	252	4	AAU07497	Aau07497 Synthetic
16	924.5	72.7	291	8	ADN06993	Adn06993 Human EFG
17	919	72.2	294	4	AAB20442	Aab20442 Anti-FIX/
18	919	72.2	732	4	AAB20437	Aab20437 Anti-FIX/
19	915	71.9	325	4	AAB20438	Aab20438 Anti-FIX/
20	909.5	71.5	239	2	AAR79866	Aar79866 Anti-EGFR
21	909	71.5	243	5	ABP45934	Abp45934 Human BLy
22	909	71.5	243	7	ADG96761	Adg96761 Single ch
23	909	71.5	249	4	AAB20435	Aab20435 Anti-FIX/
24	904	71.1	638	8	ADL27332	Adl27332 Amino aci
25	904	71.1	638	8	ADL27333	Adl27333 Amino aci
26	901	70.8	244	2	AAR79867	Aar79867 Anti-EGFR
27	900	70.8	249	4	AAB20434	Aab20434 Anti-FIX/
28	897	70.5	245	8	ADI58105	Adi58105 Reg IV-sp
29	892.5	70.2	239	2	AAY17963	Aay17963 Mouse scF
30	892	70.1	239	5	ABP45911	Abp45911 Human BLy
31	892	70.1	239	7	ADG96738	Adg96738 Single ch
32	891.5	70.1	248	5	ABP45410	Abp45410 Human BLy
33	891.5	70.1	248	7	ADG96237	Adg96237 Single ch
34	891.5	70.1	291	8	ADN06992	Adn06992 Human EFG
35	890	70.0	282	2	AAW09818	Aaw09818 VH4715-li
36	890	70.0	282	2	AAW35564	Aaw35564 HindIII-E
37	889.5	69.9	240	7	ADG30465	Adg30465 Human GMB
38	889	69.9	241	5	ABP45878	Abp45878 Human BLy
39	889	69.9	241	7	ADG96705	Adg96705 Single ch
40	888	69.8	239	2	AAR64812	Aar64812 ScFv anti
41	888	69.8	242	2	AAR79870	Aar79870 Anti-EGFR
42	888	69.8	242	2	AAR79872	Aar79872 Anti-EGFR
43	887.5	69.8	242	3	AAY58236	Aay58236 Internali
44	887	69.7	244	2	AAR79868	Aar79868 Anti-EGFR
45	886	69.7	243	5	ABP45924	Abp45924 Human BLy

ALIGNMENTS

RESULT 1
AAE38657
ID AAE38657 standard; protein; 237 AA.
XX
AC AAE38657;
XX
DT 04-DEC-2003 (first entry)
XX
DE Mouse G1 single chain Fv-recombinant antibody.
XX
KW Major histocompatibility complex; MHC; HLA-restricted antigen; cancer;
KW viral infection; autoimmune disease; gene therapy; cytostatic; virucide;
KW immunomodulator; mouse.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Domain 1. .118
FT /note= "VH domain"
FT Region 31. .35
FT /note= "Complementarity determining region"
FT Region 50. .66
FT /note= "Complementarity determining region"
FT Region 99. .108
FT /note= "Complementarity determining region"
FT Region 119. .234
FT /note= "Peptide linker"
FT Domain 135. .237
FT /note= "VL domain"
FT Region 158. .167
FT /note= "Complementarity determining region"
FT Region 182. .189
FT /note= "Complementarity determining region"
FT Region 222. .230
FT /note= "Complementarity determining region"
XX
PN WO2003068201-A2.
XX
PD 21-AUG-2003.
XX
PF 11-FEB-2003; 2003WO-IL000105.
XX
PR 13-FEB-2002; 2002US-00073301.
XX
PA (TECR) TECHNION RES & DEV FOUND LTD.
XX
PI Reiter Y, Denkberg G;
XX

DR WPI; 2003-689603/65.
DR N-PSDB; AAD58607.
XX
PT New isolated molecule comprising an antibody that binds with a human
PT major histocompatibility complex (MHC) class I being complexed with a HLA
PT -restricted antigen, useful for treating cancer, viral infection or
PT autoimmune disease.
XX
PS Claim 63; Fig 3a; 81pp; English.
XX
CC The invention relates to an isolated molecule comprising an antibody
CC specifically bindable with a binding affinity below 20 nanomolar to a
CC human major histocompatibility complex (MHC) class I being complexed with
CC a HLA-restricted antigen. The molecules, antibodies, and methods are
CC useful for treating cancer, viral infection and an autoimmune disease.
CC The invention is useful in gene therapy. The present sequence is mouse G1
CC single chain Fv-recombinant antibody
XX
SQ Sequence 237 AA;

Query Match 100.0%; Score 1272; DB 7; Length 237;
Best Local Similarity 100.0%; Pred. No. 2.2e-83;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
Db 1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60

Qy 61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
Db 61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120

Qy 121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRLL 180
Db 121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRLL 180

Qy 181 IYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
Db 181 IYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237

RESULT 2
AAE33333
ID AAE33333 standard; protein; 281 AA.
XX
AC AAE33333;
XX
DT 02-APR-2003 (first entry)
XX
DE 12E antibody.
XX
KW MUC-1; 12E; variable light domain; VL; variable heavy domain; VH;
KW diabody; cancer; antibody; therapy.
XX
OS Unidentified.
XX
PN WO200279429-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009735.
XX
PR 30-MAR-2001; 2001US-0280721P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Denardo SJ, Winthrop MD, Denardo GL;
XX
DR WPI; 2003-046804/04.
DR N-PSDB; AAD50925.
XX
PT Novel antibody that specifically binds to cancer antigen MUC-1 useful for
PT detecting a cell bearing MUC-1 antigen, comprises variable light or
PT variable heavy domains of antibodies 12E, 3D, A5 or C4.
XX
PS Claim 67; Page 16; 75pp; English.
XX
CC The invention relates to a novel antibody that specifically binds to the
CC cancer antigen MUC-1. The antibody comprises a domain having a sequence
CC of a polypeptide selected from 12E variable light (VL) or variable heavy
CC (VH) domain, 3D VL or VH domain, A5 VL or VH domain and C4 VL or VH
CC domain. Antibodies of the invention are useful for detecting a cell
CC bearing a MUC-1 antigen. The invention is useful for producing a variety
CC of human or humanised antibodies or diabodies. The invention is also
CC useful for treating cancer. The present sequence is 12E antibody
XX
SQ Sequence 281 AA;

Query Match 77.4%; Score 984; DB 6; Length 281;
Best Local Similarity 77.2%; Pred. No. 1.2e-62;
Matches 183; Conservative 24; Mismatches 30; Indels 0; Gaps 0;

Qy 1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
Db 23 QVKLQQSGTEVVKPGASVKLSCKASGYIFTSYDIDWVRQTPEQGLEWIGWIFPGEGSTEY 82

Qy 61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
Db 83 NEKFKGRATLSVDKSSSTAYMELTRLTSEDSAVYFCARGDYYRRYFDLWGQGTTVTVSSR 142

Qy 121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRLL 180
Db 143 GGGSGGGGSGGGGSDIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRLL 202

Qy 181 IYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
Db 203 IYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 259

RESULT 3
AAB11398
ID AAB11398 standard; protein; 255 AA.
XX
AC AAB11398;
XX
DT 22-FEB-2001 (first entry)
XX
DE E. coli expression plasmid pUBS520-ScFvOx encoded protein.
XX
KW Eukaryotic protein; protease; interferon; antibody; hormone;
KW disulfide bridge.
XX
OS Escherichia coli.
OS Synthetic.
XX
PN EP1048732-A1.
XX
PD 02-NOV-2000.
XX
PF 26-APR-1999; 99EP-00107412.
XX
PR 26-APR-1999; 99EP-00107412.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
DR WPI; 2000-674185/66.
DR N-PSDB; AAC66074.
XX
PT Preparation of water-soluble eukaryotic polypeptides with disulfide
PT bridges e.g. rPA, comprises cultivation of prokaryotic cells in the
PT presence of arginine or amide compound.
XX
PS Example 6; Page 22-23; 40pp; German.
XX
CC This invention describes a novel preparation of a water-soluble,
CC naturally occurring eukaryotic polypeptide containing two or more

```
CC   cysteine units bound via a disulfide bridge which comprises cultivation
CC   of prokaryotic cells in the presence of arginine or an amide compound.
CC   The method is useful for the preparation of eukaryotic proteins e.g.
CC   proteases, interferons, protein hormones, antibodies or antibody
CC   fragments (e.g. a single chain FV fragment that binds to thyroid
CC   stimulating hormone). It is especially useful for preparing proteins with
CC   more than five disulfide bridges, e.g. recombinant plasminogen activator
CC   (rPA). The technique is simple and does not require in vitro after-
CC   treatment, such as the removal of inclusion bodies, reduction or
CC   naturization
XX
SQ   Sequence 255 AA;

  Query Match             76.8%;  Score 977.5;  DB 3;  Length 255;
  Best Local Similarity   78.9%;  Pred. No. 3e-62;
  Matches  187;  Conservative   19;  Mismatches   26;  Indels    5;  Gaps    1;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              :|||||||||||:|||| ||||||||||||:|| |||| |:| ||||| |||| |  ||
Db          3 EVKLQESGGGLVQPGGSRKLSCAASGFTFSSFGMHWVRQAPEKGLEWVAYISSGSSTIYY 62

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
               |:||||||||||| ||||:|||:||:||||||||||      |   ||||||||||||
Db         63 ADTVKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCARD-----YGAYWGQGTTVTVSSG 117

Qy        121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRLL 180
              |||||||||||||:|||||||||||||:||||||||||:||: |:||| |:||:
Db        118 GGGSGGGGSGGGGSDIELTQSPAIMSASPGEKVTMTCSASSSVRYMNWFQQKSGTSPKRW 177

Qy        181 IYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
              ||||| :: ||| |||||||||||||||: ||||||||||||:||  | ||| ||||
Db        178 IYDTSKLSSGVPARFSGSGSGTSYSLTISSMEAEDAATYYCQQWSSNPLTFGAGTKL 234


RESULT 4
AAB74199
ID   AAB74199 standard; protein; 255 AA.
XX
AC   AAB74199;
XX
DT   29-MAY-2001  (first entry)
XX
DE   PelB-scFvOxazolon fusion protein.
XX
KW   Molecular chaperone; PelB signal sequence; scFvOxazolon.
XX
OS   Unidentified.
XX
PN   EP1077262-A1.
XX
PD   21-FEB-2001.
XX
PF   24-JUL-2000; 2000EP-00115839.
XX
PR   29-JUL-1999;   99EP-00114811.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX
PI   Ambrosius D,  Rudolph R,  Schaeffner J,  Schwarz E;
XX
DR   WPI; 2001-246712/26.
DR   N-PSDB; AAF77806.
XX
PT   Producing naturally folded eukaryotic proteins e.g. antibodies,
PT   interferon, hormones or proteases that contain two or several cysteines
PT   linked by disulfide bridges comprises co-expression of a molecular
PT   chaperone.
XX
PS   Disclosure; Page 19; 35pp; English.
XX
CC   The present invention relates to a method for production of a naturally
CC   folded eukaryotic protein containing two or more cysteines linked by
CC   disulfide bridges. The method comprises co-expression and secretion into
CC   the periplasm of a molecular chaperone via an expression vector coding
CC   for the chaperone. The expression vector also encodes a signal sequence.
CC   The method is useful for producing a naturally folded eukaryotic protein
CC   such as an antibody, antibody fragment, interferon, protein hormone or a
CC   protease containing two or several cysteines linked by disulfide bridges.
CC   The present sequence is a fusion protein composed of the PelB signal
CC   sequence and ScFvOxazolon. This sequence was used in the method of the
CC   present invention
XX
SQ   Sequence 255 AA;

  Query Match             76.8%;  Score 977.5;  DB 4;  Length 255;
  Best Local Similarity   78.9%;  Pred. No. 3e-62;
  Matches  187;  Conservative   19;  Mismatches   26;  Indels    5;  Gaps    1;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              :|||||||||||:|||| ||||||||||||:|| |||| |:| ||||| |||| |  ||
Db          3 EVKLQESGGGLVQPGGSRKLSCAASGFTFSSFGMHWVRQAPEKGLEWVAYISSGSSTIYY 62

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
               |:||||||||||| ||||:|||:||:||||||||||      |   ||||||||||||
Db         63 ADTVKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCARD-----YGAYWGQGTTVTVSSG 117

Qy        121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRLL 180
              |||||||||||||:|||||||||||||:||||||||||:||: |:||| |:||:
Db        118 GGGSGGGGSGGGGSDIELTQSPAIMSASPGEKVTMTCSASSSVRYMNWFQQKSGTSPKRW 177

Qy        181 IYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
              ||||| :: ||| |||||||||||||||: ||||||||||||:||  | ||| ||||
Db        178 IYDTSKLSSGVPARFSGSGSGTSYSLTISSMEAEDAATYYCQQWSSNPLTFGAGTKL 234


RESULT 5
AAB70769
ID   AAB70769 standard; protein; 255 AA.
XX
AC   AAB70769;
XX
DT   18-MAY-2001  (first entry)
XX
DE   Expression plasmid pUBS520-ScFvOx protein.
XX
KW   Chaperone protein; periplasm; antibody production; protein production;
KW   interferon production; protease production.
XX
OS   Escherichia coli.
OS   Synthetic.
XX
PN   EP1077263-A1.
XX
PD   21-FEB-2001.
XX
PF   29-JUL-1999;   99EP-00114811.
XX
PR   29-JUL-1999;   99EP-00114811.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX
DR   WPI; 2001-204356/21.
DR   N-PSDB; AAF61193.
XX
PT   Preparation of naturally folded eukaryotic proteins, e.g. antibodies, by
PT   simultaneous expression of a chaperone protein, allows simple recovery
PT   from periplasm or medium.
XX
PS   Disclosure; Page 20-21; 36pp; German.
XX
CC   This invention describes a novel method for preparing a naturally folded
CC   eukaryotic polypeptide (I) that contains two or more disulfide-bridged
CC   Cys residues by culturing prokaryotic cells that contain an expression
CC   vector for (I) including a prokaryotic signal sequence at its N-terminus
CC   and a nucleic acid (II) that secretes a chaperone protein (III) into the
CC   periplasm. (I) is secreted into the periplasm or medium; the signal
```

CC peptide is then cleaved and (I) isolated from the periplasm or medium.
CC The method is used for production of antibody, interferon, protein
CC hormone or protease. Expression of (III) increases the yield of (I). The
CC method is simple and eliminates time-consuming in vitro processing
CC operations such as dissolution of inclusion bodies, reduction and
CC refolding. (III) protects (I) against agglomeration and promotes their
CC natural conformation
XX
SQ Sequence 255 AA;

Query Match 76.8%; Score 977.5; DB 4; Length 255;
Best Local Similarity 78.9%; Pred. No. 3e-62;
Matches 187; Conservative 19; Mismatches 26; Indels 5; Gaps 1;

```
Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              :|||||||||||:|||| ||||||||||||:|| |||| |:| ||||| |||| |  ||
Db          3 EVKLQESGGGLVQPGGSRKLSCAASGFTFSSFGMHWVRQAPEKGLEWVAYISSGSSTIYY 62

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
               |:|||||||||| |||:|||:||:||||||||||      |   ||||||||||
Db         63 ADTVKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCARD-----YGAYWGQGTTVTVSSG 117

Qy        121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRLL 180
              |||||||||||||:|||||||||||||:||||||||||:||: |:||| |:||:
Db        118 GGGSGGGGSGGGGSDIELTQSPAIMSASPGEKVTMTCSASSSVRYMNWFQQKSGTSPKRW 177

Qy        181 IYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
              ||||| :: ||| |||||||||||||||: ||||||||||||:||  | ||| ||||
Db        178 IYDTSKLSSGVPARFSGSGSGTSYSLTISSMEAEDAATYYCQQWSSNPLTFGAGTKL 234
```

RESULT 6
AAY72020
ID AAY72020 standard; protein; 255 AA.
XX
AC AAY72020;
XX
DT 28-MAR-2001 (first entry)
XX
DE E. carotovora PelB-scFvOx fusion protein encoded by pUBS520-pIN-scFvOx.
XX
KW Secreted protein; chaperone; interferon; protease; hormone;
KW naturally folded protein; lac promoter; DnaJ; heat shock protein; HSP;
KW pectate lyase B; PelB; hapten; single-chain Fv-fragment Oxazolon;
KW scFvOxazolon; fusion protein; thyroid stimulating hormone; TSH.
XX
OS Pectobacterium carotovorum.
OS Unidentified.
OS Chimeric.
XX
PN EP1054063-A2.
XX
PD 22-NOV-2000.
XX
PF 19-APR-2000; 2000EP-00108505.
XX
PR 26-APR-1999; 99EP-00107412.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
PI Ambrosius D, Rudolph R, Schaeffner J, Schwarz E;
XX
DR WPI; 2001-033777/05.
DR N-PSDB; AAD02212.
XX
PT Producing water-soluble, naturally folded, and secreted eukaryotic
PT polypeptide, involves culturing prokaryotic cells containing an
PT expression vector encoding the polypeptide in the presence of arginine or
PT a specific compound.
XX
PS Example 6; Page 22-23; 35pp; English.
XX
CC The patent discloses a method for the production of a water-soluble,
CC naturally folded and secreted eukaryotic proteins in prokaryotic cells.
CC The method involves culturing the prokaryotic cells, containing an
CC expression vector encoding the desired protein and the prokaryotic signal
CC sequence, in the presence of an additive, e.g., L-arginine. The signal
CC sequence mediates the secretion of the desired protein into the
CC periplasm, where folding of the protein takes place. The prokaryotic cell
CC also contains an expression vector encoding a molecular chaperone, e.g.,
CC DnaJ and heat shock protein 25 (HSP25). The simultaneous co-
CC overexpression and co-secretion of molecular chaperones in the periplasm
CC improves the yield of functionally folded protein. The above method
CC recombinantly produces a high yield of eukaryotic secreted proteins in
CC prokaryotes. The method is useful for producing eukaryotic proteins such
CC as an antibody, antibody fragment, interferon, protein hormone or a
CC protease. The present sequence is an Erwinia carotovora pectate lyase B
CC (PelB) signal sequence-scFvOx fusion protein encoded by pUBS520-pIN-
CC scFvOx. The plasmid, pUBS520-pIN-scFvOx, also comprises the lac promoter.
CC The single-chain Fv-fragment, which is directed against the hapten
CC oxazolon (ScFvOxazolon), is an antibody fragment against thyroid
CC stimulating hormone (TSH). The co-expression of ScFvOx which has no
CC chaperone properties is used as a negative control
XX
SQ Sequence 255 AA;

Query Match 76.8%; Score 977.5; DB 4; Length 255;
Best Local Similarity 78.9%; Pred. No. 3e-62;
Matches 187; Conservative 19; Mismatches 26; Indels 5; Gaps 1;

```
Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              :|||||||||||:|||| ||||||||||||:|| |||| |:| ||||| |||| |  ||
Db          3 EVKLQESGGGLVQPGGSRKLSCAASGFTFSSFGMHWVRQAPEKGLEWVAYISSGSSTIYY 62

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
               |:|||||||||| |||:|||:||:||||||||||      |   ||||||||||
Db         63 ADTVKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCARD-----YGAYWGQGTTVTVSSG 117

Qy        121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRLL 180
              |||||||||||||:|||||||||||||:||||||||||:||: |:||| |:||:
Db        118 GGGSGGGGSGGGGSDIELTQSPAIMSASPGEKVTMTCSASSSVRYMNWFQQKSGTSPKRW 177

Qy        181 IYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
              ||||| :: ||| |||||||||||||||: ||||||||||||:||  | ||| ||||
Db        178 IYDTSKLSSGVPARFSGSGSGTSYSLTISSMEAEDAATYYCQQWSSNPLTFGAGTKL 234
```

RESULT 7
AAY32086
ID AAY32086 standard; protein; 316 AA.
XX
AC AAY32086;
XX
DT 17-JAN-2000 (first entry)
XX
DE Single chain antibody used in probe detection.
XX
KW Single chain antibody; scAb; sFv; spectroscopic probe.
XX
OS Unidentified.
XX
PN WO9951986-A1.
XX
PD 14-OCT-1999.
XX
PF 08-APR-1999; 99WO-US007847.
XX
PR 08-APR-1998; 98US-0081118P.
PR 09-APR-1998; 98US-0081340P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Farinas J;
XX
DR WPI; 1999-611066/52.
DR N-PSDB; AAZ20266.

XX
PT Localizing probes to specific sites in cells that express single-chain
PT antibody reactive with probe-ligand conjugate, particularly for detecting
PT post-translational modification and its modulators.
XX
PS Claim 5; Page 47-48; 69pp; English.
XX
CC The present sequence represents a single chain antibody (scAb) that has 2
CC c-myc epitopes. The scAb is used in claimed methods for localizing a
CC probe and for detecting a post-translational activity in a cell, and is
CC expressed by a claimed transgenic non-human animal. The invention
CC provides methods and reagents for targeting probes to selected cellular
CC locations through the expression of specific binding partners (such as
CC scAbs) within a cell, and for creating assays for post-translational
CC activities. The invention allows the monitoring of the location of such
CC intracellular, specific binding patrtners over time and in response to
CC stimuli, such as test chemicals. Spectroscopic probes can be used to
CC screen a test chemical for activity. The invention also includes cells
CC and transgenic organisms comprising the intracellular specific binding
CC partner, where the specific binding partner can bind with the
CC spectroscopic probe/ligand conjugate. scAbs can be expressed within the
CC cell and can be designed to bind a wide variety of spectroscopic probes,
CC including small molecules that have better (and more diverse)
CC spectroscopic properties than green fluorescent protein
XX
SQ Sequence 316 AA;

Query Match 76.8%; Score 976.5; DB 2; Length 316;
Best Local Similarity 78.9%; Pred. No. 4.5e-62;
Matches 187; Conservative 18; Mismatches 27; Indels 5; Gaps 1;

```
Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              :|||||||||||:|||| |||||||||||||:|| |||| |:| ||||| |||| |  ||
Db          3 EVKLQESGGGLVQPGGSRKLSCAASGFTFSSFGMHWVRQAPEKGLEWVAYISSGSSTIYY 62

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
               |:||||||||||| ||||:|||:||:|||| ||||||      |   ||||||||||||
Db         63 ADTVKGRFTISRDNPKNTLFLQMTSLRSEDTVMYYCARD-----YGAYWGQGTTVTVSSG 117

Qy        121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRLL 180
              |||||||||||||:|||||||||||||||||||||||:||: |:||| |:||:
Db        118 GGGSGGGGSGGGGSDIELTQSPAIMSASPGERVTMTCSASSSVRYMNWFQQKSGTSPKRW 177

Qy        181 IYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
              ||||| :: ||| |||||||||||||||: ||||||||||:||  | ||| ||||
Db        178 IYDTSKLSSGVPARFSGSGSGTSYSLTISSMEAEDAATYYCQQWSSNPLTFGAGTKL 234
```

RESULT 8
ABR62010
ID ABR62010 standard; protein; 241 AA.
XX
AC ABR62010;
XX
DT 03-OCT-2003 (first entry)
XX
DE Single-chain (scFv) antibody.
XX
KW Transgenic; biomolecule; ATP; ADP; cytostatic; virucide; gene therapy;
KW antisense therapy; scFv; antibody.
XX
OS Synthetic.
XX
PN EP1321524-A1.
XX
PD 25-JUN-2003.
XX
PF 19-DEC-2001; 2001EP-00130319.
XX
PR 19-DEC-2001; 2001EP-00130319.
XX
PA (DUER/) DUERING K.
XX
PI Mahn A, Hantke S, Petsch D;
XX
DR WPI; 2003-543829/52.
DR N-PSDB; ACC84876.
XX
PT Increasing the content of transgene-coded biomolecules in a plant or
PT animal, useful for producing proteins for diagnosing, preventing and/or
PT treating viral diseases and cancer, comprises changing the distribution
PT of ATP and/or ADP.
XX
PS Example 2; Fig 2; 18pp; English.
XX
CC The invention relates to increasing the content of one or more transgene-
CC coded biomolecules in an organism and involves changing the distribution
CC of ATP and/or ADP in cells of the organism. The yield of transgenic
CC molecules in host cells is often insufficient for industrial production.
CC The method increases the yield of transgenic molecules in animal and
CC plant host cells, therefore facilitating their production on an
CC industrial scale. The proteins produced by the method are useful for
CC diagnosing, preventing and/or treating viral diseases and cancer. The
CC present sequence represents a single-chain (scFv) antibody, used to
CC exemplify the increase in the expression of scFv antibodies in transgenic
CC potato tubers
XX
SQ Sequence 241 AA;

Query Match 75.8%; Score 964.5; DB 6; Length 241;
Best Local Similarity 78.4%; Pred. No. 2.4e-61;
Matches 185; Conservative 19; Mismatches 27; Indels 5; Gaps 1;

```
Qy          2 VKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYYP 61
              |:| |||||||:|||| ||||||||||||:|| |||| |:| ||||| |||| |  ||
Db          4 VQLVESGGGLVQPGGSRKLSCAASGFTFSSFGMHWVRQAPEKGLEWVAYISSGSSTIYYA 63

Qy         62 DSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSGG 121
              |:||||||||||| ||||:|||:||:|||||||||||      |   ||||||||||||
Db         64 DTVKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCARD-----YGAYWGQGTTVTVSSGG 118

Qy        122 GGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRLLI 181
              |||||||||||||:||||||||||||||||:||||||||||:||: |:||| |:||:  |
Db        119 GGSGGGGSGGGGSDIELTQSPAIMSASPGEKVTMTCSASSSVRYMNWFQQKSGTSPKRWI 178

Qy        182 YDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
              |||| :: ||| |||||||||||||||: ||||||||||:||  | ||| ||||
Db        179 YDTSKLSSGVPARFSGSGSGTSYSLTISSMEAEDAATYYCQQWSSNPLTFGAGTKL 234
```

RESULT 9
AAM48925
ID AAM48925 standard; protein; 241 AA.
XX
AC AAM48925;
XX
DT 03-MAY-2002 (first entry)
XX
DE scFv antibody.
XX
KW Ketone binding protein; oxazole; pathogen resistance; virucide;
KW fungicide; antibacterial; scFv antibody.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 137
FT /note= "encoded by CAG"
FT Misc-difference 201
FT /note= "encoded by TAC"
XX
PN WO200204020-A2.
XX
PD 17-JAN-2002.
XX
PF 17-MAY-2001; 2001WO-DE001916.

```
XX
PR   12-JUL-2000; 2000DE-01033750.
XX
PA   (MPBC-) MPB COLOGNE GMBH.
XX
PI   Duering K,  Brinkmann O;
XX
DR   WPI; 2002-154868/20.
DR   N-PSDB; AAK98639.
XX
PT   Imparting pathogen resistance to plants and animals, comprises using a
PT   polypeptide that binds to an oxazole-derived ketone, optionally expressed
PT   from nucleic acid.
XX
PS   Example 3; Fig 1; 20pp; German.
XX
CC   The present invention relates to the use of a polypeptide that binds a
CC   ketone derived from oxazole for generating pathogen resistance in an
CC   organism. This can be used to impart resistance to pathogens (bacteria,
CC   fungi or viruses) to a wide range of plants (e.g. cereals, sugar beet,
CC   tobacco etc.), humans, farm animals and pets. Exemplified are activities
CC   against Phytophthora infestans and Erwinia carotovora in potatoes and
CC   Staphylococcus aureus in mice. The present sequence is a sc-Fv antibody,
CC   which is capable of binding 2-phenyloxazol-5-one
XX
SQ   Sequence 241 AA;

  Query Match             74.6%;  Score 949.5;  DB 5;  Length 241;
  Best Local Similarity   77.5%;  Pred. No. 2.9e-60;
  Matches  183;  Conservative   19;  Mismatches   29;  Indels    5;  Gaps    1;

Qy          2 VKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYYP 61
              |:| |||||||:|||| |||||||||||||:|| |||| |:| ||||| |||| |  ||
Db          4 VQLVESGGGLVQPGGSRKLSCAASGFTFSSFGMHWVRQAPEKGLEWVAYISSGSSTIYYA 63

Qy         62 DSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSGG 121
              |:||||||||||| ||||:|||:||:|||||||||||      |   ||||||||||||
Db         64 DTVKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCARD-----YGAYWGQGTTVTVSSGG 118

Qy        122 GGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRLLI 181
              |||||||||||||:|||| ||||||||||:||||||||||:||: |:||| |:||:  |
Db        119 GGSGGGGSGGGGSDIELTGSPAIMSASPGEKVTMTCSASSSVRYMNWFQQKSGTSPKRWI 178

Qy        182 YDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
              |||| :: ||| |||||||||| ||||: ||||||||||||:||  | ||| ||||
Db        179 YDTSKLSSGVPARFSGSGSGTSVSLTISSMEAEDAATYYCQQWSSNPLTFGAGTKL 234


RESULT 10
AAR32842
ID   AAR32842 standard; protein; 236 AA.
XX
AC   AAR32842;
XX
DT   25-MAR-2003  (revised)
DT   19-JUN-1993  (first entry)
XX
DE   VH NQ10/12.5-Vk NQ10/12.5 linked peptide sequences #2.
XX
KW   Primer; human; immunoglobulin; Ig; variable region; VH; VL; Ck; JH;
KW   lymphocyte; vector; soluble; antibody; phage; linker; back; VH3; nested;
KW   in-cell PCR; cloning; polymorphic; TCR V; antiphenyloxazolone; hybridoma;
KW   NQ2/12.4; NQ10/12.5.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .115
FT                   /label= VH_NQ10/12.5
FT   Peptide         116. .129
FT                   /note= "Linker peptide"
FT   Region          130. .236
FT                   /label= Vkappa_NQ10/12.5
XX
PN   WO9303151-A1.
XX
PD   18-FEB-1993.
XX
PF   10-AUG-1992;   92WO-GB001483.
XX
PR   10-AUG-1991;   91GB-00017352.
PR   11-JUN-1992;   92GB-00012419.
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
XX
PI   Embleton MJ,  Gorochov G,  Jones PT,  Winter GP;
XX
DR   WPI; 1993-076508/09.
DR   N-PSDB; AAQ37461.
XX
PT   Treatment of cell populations, partic. hybridomas - to link together
PT   copies of 2 or more non-contiguous DNA sequences to facilitate analysis.
XX
PS   Disclosure; Fig 4; 72pp; English.
XX
CC   The sequences given in AAR32840-43 show the mature heavy chain VH domains
CC   and the Vk light chain genes of the antiphenyloxazolone hybridomas
CC   NQ2/12.4 and NQ10/12.5 which have been linked via a linker peptide by in-
CC   cell PCR. The cDNA encoding these peptides was synthesised using forward
CC   primers annealing to the Ck gene and the JH segment, followed by assembly
CC   with linker primers, VH back primers based on the VH3 leader sequence and
CC   a forward Ck primer nested in respect to the primer used for cDNA. The
CC   assembled product within the cells is then amplified with nested primers
CC   annealing to the 5' end of the VH gene and the 3' end of the Jk segment.
CC   In-cell PCR may be used to determine gene linkage analysis, particularly
CC   for the cloning of gene combinations that are polymorphic within a
CC   population of cells, such as the rearranged genes for Ig or TCR V
CC   regions. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 236 AA;

  Query Match             74.5%;  Score 947.5;  DB 2;  Length 236;
  Best Local Similarity   77.1%;  Pred. No. 4e-60;
  Matches  182;  Conservative   19;  Mismatches   30;  Indels    5;  Gaps    1;

Qy          2 VKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYYP 61
              |:| |||||||:|||| |||||||||||||:|| |||| |:| ||||| |||| |  ||
Db          2 VQLVESGGGLVQPGGSRKLSCAASGFTFSSFGMHWVRQAPEKGLEWVAYISSGSSTIYYA 61

Qy         62 DSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSGG 121
              |:||||||||||| ||||:|||:||:|||||||||||      |   ||||| ||||:||
Db         62 DTVKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCARD-----YGAYWGQGTLVTVSAGG 116

Qy        122 GGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRLLI 181
              |||||||||||||| | |||||||||||||:||||||||||:||: |:||| |:||:  |
Db        117 GGSGGGGSGGGGSQIVLTQSPAIMSASPGEKVTMTCSASSSVRYMNWFQQKSGTSPKRWI 176

Qy        182 YDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
              |||| :: ||| |||||||||||||||: ||||||||||||:||  | ||| ||||
Db        177 YDTSKLSSGVPARFSGSGSGTSYSLTISSMEAEDAATYYCQQWSSNPLTFGAGTKL 232


RESULT 11
AAB20436
ID   AAB20436 standard; protein; 249 AA.
XX
AC   AAB20436;
XX
DT   21-JUN-2001  (first entry)
XX
DE   Anti-FIX/FIXa antibody 198/A1 scFv.
XX
KW   Factor IX; FIX; Factor IXa; FIXa; scFv; antibody; procoagulant;
KW   Factor VIII cofactor; blood coagulation disorder; haemophilia A;
KW   haemorrhagic diathesis; haemostatic; amidolytic; therapy; mouse.
XX
```

```
OS   Mus musculus.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Protein         1. .122
FT                   /label= VH
FT   Region          99. .111
FT                   /label= CDR3
FT   Peptide         123. .136
FT                   /label= Linker
FT   Protein         137. .249
FT                   /label= VL
FT   Misc-difference 142
FT                   /note= "encoded by ACN"
FT   Misc-difference 224
FT                   /note= "encoded by GCN"
FT   Region          230. .238
FT                   /label= CDR3
XX
PN   WO200119992-A2.
XX
PD   22-MAR-2001.
XX
PF   13-SEP-2000; 2000WO-EP008936.
XX
PR   14-SEP-1999;   99AT-00001576.
XX
PA   (BAXT ) BAXTER AG.
XX
PI   Scheiflinger F,  Kerschbaumer R,  Falkner F,  Dorner F;
XX
DR   WPI; 2001-290358/30.
DR   N-PSDB; AAF30726.
XX
PT   New factor IX/factor IXa antibodies and their derivatives useful for
PT   increasing amidolytic activity of factor IXa, and for treating blood
PT   coagulation disorders such as hemophilia A and hemorrhagic diathesis.
XX
PS   Example 10; Fig 17; 138pp; English.
XX
CC   The present sequence is that of a single chain Fv (scFv) derivative of
CC   antibody 198/A1, comprising the heavy (VH) and light (VL) chain variable
CC   regions of 198/A1 joined by an artificial, flexible linker peptide. The
CC   scFv was obtained by PCR amplification of cDNAs for 198/A1 VH and VL
CC   regions and cloning in vector pDAP2. 198/A1 is an example of anti-human
CC   Factor IX (FIX)/activated Factor IX (FIXa) antibodies of the invention.
CC   Anti-FIX/FIXa and their derivatives, including scFv and CDR3 fragments,
CC   have Factor VIIIa (FVIIIa) cofactor activity or FIXa activating activity.
CC   Administration leads to an increase in the procoagulant activity of FIXa,
CC   even in the presence of FVIIIa inhibitors. This allows for rapid blood
CC   coagulation even in the absence of FVIII or FVIIIa, and in the case of
CC   FVIII inhibitor patients. The antibodies and derivatives are used in a
CC   claimed pharmaceutical composition for treating patients with blood
CC   coagulation disorders, especially haemophilia A and haemorrhagic
CC   diathesis
XX
SQ   Sequence 249 AA;

  Query Match             73.6%;  Score 936;  DB 4;  Length 249;
  Best Local Similarity   72.7%;  Pred. No. 2.8e-59;
  Matches  178;  Conservative   26;  Mismatches   33;  Indels    8;  Gaps    2;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              :|:|||||||||||||||||||||||| |||| ||||||||:|||||||||||||| |||
Db          1 EVQLQESGGGLVKPGGSLKLSCAASGFIFSSYTMSWVRQTPEKRLEWVATISSGGSSTYY 60

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCAR---GNWEGWYFDVWGQGTTVTV 117
              ||||||||||||||||||||||||||||||:| | ·|:: |||||||| |||:||
Db         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYHCTREGGGYYVNWYFDVWGAGTTLTV 120

Qy        118 SSGGGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSI-----RYIYWYQQK 172
              |||||||||  ||||||:|||||||| :: | |:| |::| || |:      :::||||||

Db        121 SSGGGGSGGRASGGGGSDIELTQSPASLAVSLGQRATISCRASESVDSYGKSFMHWYQQK 180

Qy        173 PGSSPRLLIYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFG 232
              ||  |:|||||  ||:  |:| ||||||| | ::|||| :||:| ||||||: :  | |||
Db        181 PGQPPKLLIYRASNLESGIPARFSGSGSRTDFTLTINPVEADDVATYYCQQSNEDPLTFG 240

Qy        233 GGTKL 237
                ||:|
Db        241 AGTRL 245
```

```
RESULT 12
AAR68613
ID   AAR68613 standard; protein; 240 AA.
XX
AC   AAR68613;
XX
DT   25-MAR-2003  (revised)
DT   13-SEP-1995  (first entry)
XX
DE   Single chain antibody (scFV) which binds to phenyloxazolone.
XX
KW   Genetic selection; ligand binding protein; cholera toxin; promoter;
KW   detection; selection; beta galactosidase; lac; transmembrane domain;
KW   regulatory domain; ds.
XX
OS   Synthetic.
XX
PN   DE4319296-A1.
XX
PD   15-DEC-1994.
XX
PF   10-JUN-1993;   93DE-04319296.
XX
PR   10-JUN-1993;   93DE-04319296.
XX
PA   (BEHW ) BEHRINGWERKE AG.
XX
PI   Fritz H,  Hennecke F,  Kolmar H;
XX
DR   WPI; 1995-023689/04.
DR   N-PSDB; AAQ80468.
XX
PT   Genetic selection of ligand binding proteins in microorganisms - by
PT   extracytoplasmic protein presentation, then use of ligand binding to
PT   express a detectable or selectable function.
XX
PS   Example 2.2; Fig 4; 25pp; German.
XX
CC   Genetic selection in microorganisms, for ligand binding proteins (LBP)
CC   comprises: extracytoplasmic presentation of LBP and; using the signal of
CC   ligand binding (by signal transduction on the biosynthetic machinery of
CC   the microorganisms) to express a detectable or selectable function.
CC   Microorganisms for this process include a genetically stable
CC   detection/selection and are transformed with a replicon encoding a fusion
CC   protein consisting of the LBP, a transmembrane helix and regulatory
CC   domain. The detection/selection function is expression of a beta-
CC   galactosidase gene, integrated into the chromosome and under the control
CC   of the ctx (cholera toxin) promoter. The transmembrane helix is taken
CC   from the toxR gene. Four primers (AAQ80457-60) were used in the
CC   construction of the plasmid pHKToxscFV. The primers described in AAQ80457
CC   -58 were used to amplify variable heavy chain sequence of the single
CC   chain antibody NQ10.12.5 and those described in AAQ80459-60 were used to
CC   amplify the corresponding light chain sequence from the same antibody.
CC   The amplified sequences were cloned into the plasmid pHKToxREI (See
CC   AAQ80454-56 for details) to create a toxR-scFV fusion gene. This sequence
CC   is the single chain antibody (scFV). (Updated on 25-MAR-2003 to correct
CC   PN field.)
XX
SQ   Sequence 240 AA;

  Query Match             73.5%;  Score 935.5;  DB 2;  Length 240;
  Best Local Similarity   76.3%;  Pred. No. 2.9e-59;
```

```
  Matches  180;  Conservative   18;  Mismatches   33;  Indels    5;  Gaps    1;

Qy          2 VKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYYP 61
              |:| | ||| |:|||| ||||||||||||:|| |||| |:| ||||| |||| |  ||
Db          6 VQLVELGGGFVQPGGSRKLSCAASGFTFSSFGMHWVRQAPEKGLEWVAYISSGSSTIYYA 65

Qy         62 DSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSGG 121
              |:||||||||||| |||:|||:||:|||||||||||      |   ||||| |||||||
Db         66 DTVKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCARD-----YGAYWGQGTLVTVSSGG 120

Qy        122 GGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRLLI 181
              |||||||||||| | ||||||||||||:||||||||||:||: |:||| |:||:
Db        121 GGSGGGGSGGGGSQIVLTQSPAIMSASPGEKVTMTCSASSSVRYMNWFQQKSGTSPKRWT 180

Qy        182 YDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
              |||| :: ||| |||||||||||||||: ||||||||||:||  | ||| ||||
Db        181 YDTSKLSSGVPARFSGSGSGTSYSLTISSMEAEDAATYYCQQWSSNPLTFGAGTKL 236


RESULT 13
AAW90218
ID   AAW90218 standard; protein; 556 AA.
XX
AC   AAW90218;
XX
DT   10-MAY-1999  (first entry)
XX
DE   Bispecific tetravalent antibody BiTAb1G10-B7-24H6.
XX
KW   B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2; CD86;
KW   T cell activation; inhibitor; graft versus host disease;
KW   transplant rejection; allograft rejection; autoimmune disease; allergy;
KW   therapy; human; bispecific tetravalent antibody; BiTAb;
KW   BiTAb1G10-B7-24H6.
XX
OS   Mus sp.
OS   Homo sapiens.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .120
FT                   /note= "VH region anti B7.2 MAb"
FT   Peptide         121. .135
FT                   /note= "(G4S3) flexible linker"
FT   Region          136. .248
FT                   /note= "VL region anti B7.2 MAb"
FT   Region          249. .259
FT                   /note= "human IgG3 hinge region"
FT   Domain          260. .285
FT                   /note= "helix-turn-helix dimerisation domain"
FT   Domain          286. .305
FT                   /note= "human IgG3 hinge domain"
FT   Region          306. .426
FT                   /note= "VH region anti B7.1 MAb"
FT   Peptide         427. .441
FT                   /note= "(G4S3)flexible linker"
FT   Region          442. .550
FT                   /note= "VL region anti B7.1 MAb"
FT   Peptide         551. .556
FT                   /note= "His6 tag"
XX
PN   WO9858965-A2.
XX
PD   30-DEC-1998.
XX
PF   22-JUN-1998;   98WO-EP003791.
XX
PR   20-JUN-1997;   97EP-00870092.
XX
PA   (INNO-) INNOGENETICS NV.
XX
PI   Lorre K,  Sablon E,  Buyse M,  Bosman A;
XX
DR   WPI; 1999-105615/09.
DR   N-PSDB; AAX01652.
XX
PT   New molecules which bind B7.1 and B7.2 - useful to prevent and treat
PT   immune diseases including allograft rejection.
XX
PS   Example 7.1; Fig 18; 182pp; English.
XX
CC   This polypeptide comprises the bispecific tetravalent antibody BiTAb1G10-
CC   B7-24H6. The molecule consists of 4 scFvs, i.e. 2 anti B7.1 scFvs and 2
CC   anti B7.2 scFvs (tetravalency). One single BiTAb is a homodimer of 2
CC   identical molecules, each containing both an anti B7.1 and anti B7.2 scFv
CC   (bispecificity). An anti-B7.1 and and anti-B7.2 scFv are linked using a
CC   dimerisation domain (see AAW90219), which drives the homodimerisation of
CC   the molecule. DNA (see AAX01652) encoding the BiTAb has been constructed
CC   to allow expression of the BiTAb in transformed E. coli cells. The BiTAb
CC   cross-links, and/or cross-reacts, with the costimulatory molecules B7.1
CC   and B7.2 that are expressed on the membrane of professional antigen-
CC   presenting cells, leading to the inhibition of antigen- specific T cell
CC   activation. The invention relates to such B7-binding molecules, methods
CC   for their production, and their use for treating or preventing diseases
CC   of the immune system, in particular graft rejection, graft versus host
CC   disease, allergy and autoimmune diseases (claimed)
XX
SQ   Sequence 556 AA;

  Query Match             73.3%;  Score 932.5;  DB 2;  Length 556;
  Best Local Similarity   74.5%;  Pred. No. 1.2e-58;
  Matches  178;  Conservative   25;  Mismatches   31;  Indels    5;  Gaps    2;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              ||:|||||||||:|||| :||||||||||||:|| |||| | | ||||| |||  :  ||
Db        311 QVQLQESGGGLVQPGGSRRLSCAASGFTFSSFGMHWVRQAPGKGLEWVAFISSVSTLIYY 370

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
              |||||||||||||||||:|||:||::||||:|||||   :|||||||||||| ||||
Db        371 ADSVKGRFTISRDNAKNTLFLQMNSLRAEDTAVYYCAR---DGWYFDVWGQGTTVIVSSG 427

Qy        121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSI--RYIYWYQQKPGSSPR 178
              |||||||||||||:|||||||: |:|| |:|||:||| || |    ::|||||  :||:
Db        428 GGGSGGGGSGGGGSDIELTQSPSSMAASVGDRVTITCSVSSRISSSNLHWYQQKSETSPK 487

Qy        179 LLIYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
              || |||:| ||| |||||||||| |:|||: |: |||||||||:|| || ||| ||||
Db        488 PWIYGTSNLASGVPSRFSGSGSGTDYTLTISSMQPEDAATYYCQQWSSYPLTFGQGTKL 546


RESULT 14
AAW90217
ID   AAW90217 standard; protein; 580 AA.
XX
AC   AAW90217;
XX
DT   10-MAY-1999  (first entry)
XX
DE   Bispecific tetravalent antibody BiTAbB7-24-IG10H6.
XX
KW   B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2; CD86;
KW   T cell activation; inhibitor; graft versus host disease;
KW   transplant rejection; allograft rejection; autoimmune disease; allergy;
KW   therapy; human; bispecific tetravalent antibody; BiTAb;
KW   BiTAbB7-24-1G10H6.
XX
OS   Mus sp.
OS   Homo sapiens.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .24
FT                   /note= "pelB signal peptide"
FT   Region          25. .138
```

```
FT                   /note= "VH region anti B7.1 MAb"
FT   Peptide         139. .153
FT                   /note= "(G4S3) flexible linker"
FT   Region          154. .262
FT                   /note= "VL region anti B7.1 MAb"
FT   Misc-difference 261
FT                   /note= "encoded by CTG"
FT   Region          263. .273
FT                   /note= "human IgG3 hinge region"
FT   Domain          274. .308
FT                   /note= "helix-turn-helix dimerisation domain"
FT   Domain          309. .319
FT                   /note= "human IgG3 hinge domain"
FT   Region          320. .446
FT                   /note= "VH region anti B7.2 MAb"
FT   Misc-difference 322. .327
FT                   /note= "codons for these amino acids are not present in
FT                   the DNA sequence for BTiTAbB7-24-1G1-H6 provided in the
FT                   specification"
FT   Peptide         447. .461
FT                   /note= "(G4S3)flexible linker"
FT   Region          462. .574
FT                   /note= "VL region anti B7.2 MAb"
FT   Peptide         575. .580
FT                   /note= "His6 tag"
XX
PN   WO9858965-A2.
XX
PD   30-DEC-1998.
XX
PF   22-JUN-1998;   98WO-EP003791.
XX
PR   20-JUN-1997;   97EP-00870092.
XX
PA   (INNO-) INNOGENETICS NV.
XX
PI   Lorre K,  Sablon E,  Buyse M,  Bosman A;
XX
DR   WPI; 1999-105615/09.
DR   N-PSDB; AAX01651.
XX
PT   New molecules which bind B7.1 and B7.2 - useful to prevent and treat
PT   immune diseases including allograft rejection.
XX
PS   Example 7.1; Fig 16; 182pp; English.
XX
CC   This polypeptide comprises the bispecific tetravalent antibody BiTAbB7-24
CC   -1G10H6. The molecule consists of 4 scFvs, i.e. 2 anti B7.1 scFvs and 2
CC   anti B7.2 scFvs (tetravalency). One single BiTAb is a homodimer of 2
CC   identical molecules, each containing both an anti B7.1 and anti B7.2 scFv
CC   (bispecificity). An anti-B7.1 and and anti-B7.2 scFv are linked using a
CC   dimerisation domain (see AAW90219), which drives the homodimerisation of
CC   the molecule. DNA (see AAX01651) encoding the BiTAb has been constructed
CC   to allow expression of the BiTAb in transformed E. coli cells. The BiTAb
CC   cross-links, and/or cross-reacts, with the costimulatory molecules B7.1
CC   and B7.2 that are expressed on the membrane of professional antigen-
CC   presenting cells, leading to the inhibition of antigen- specific T cell
CC   activation. The invention relates to such B7-binding molecules, methods
CC   for their production, and their use for treating or preventing diseases
CC   of the immune system, in particular graft rejection, graft versus host
CC   disease, allergy and autoimmune diseases (claimed)
XX
SQ   Sequence 580 AA;

  Query Match             73.3%;  Score 932.5;  DB 2;  Length 580;
  Best Local Similarity   74.5%;  Pred. No. 1.2e-58;
  Matches  178;  Conservative   25;  Mismatches   31;  Indels    5;  Gaps    2;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              ||:|||||||||:|||| :|||||||||||||:|| |||| | | ||||| |||  :  ||
Db         23 QVQLQESGGGLVQPGGSRRLSCAASGFTFSSFGMHWVRQAPGKGLEWVAFISSVSTLIYY 82

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
                ||||||||||||||||||:|||:||::||||:|||||   :|||||||||||| ||||
Db         83 ADSVKGRFTISRDNAKNTLFLQMNSLRAEDTAVYYCAR---DGWYFDVWGQGTTVIVSSG 139

Qy        121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSI--RYIYWYQQKPGSSPR 178
              |||||||||||||:|||||||: |:|| |:|||:||| || |    ::|||||  :||:
Db        140 GGGSGGGGSGGGGSDIELTQSPSSMAASVGDRVTITCSVSSRISSSNLHWYQQKSETSPK 199

Qy        179 LLIYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
                || |||:| ||| |||||||||| |:|||: |: |||||||||:|| || ||| ||||
Db        200 PWIYGTSNLASGVPSRFSGSGSGTDYTLTISSMQPEDAATYYCQQWSSYPLTFGQGTKL 258


RESULT 15
AAU07497
ID   AAU07497 standard; protein; 252 AA.
XX
AC   AAU07497;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Synthetic antibody scFv(F8).
XX
KW   Antimicrobial; antiviral; cytostatic; immunomodulatory; antibody;
KW   gene therapy; HIV; light chain; human immunodeficiency virus; tumour;
KW   metabolic disorder; immune disorder; auto-immune disorder; scFv(F8);
KW   cucumber mosaic virus.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Protein         1. .125
FT                   /label= VH
FT                   /note= "Heavy chain variable region"
FT   Peptide         126. .140
FT                   /label= Linker_peptide
FT                   /note= "This peptide is specifically claimed in claim 17"
FT   Protein         141. .252
FT                   /label= VL
FT                   /note= "Light chain variable region"
XX
PN   WO200149713-A2.
XX
PD   12-JUL-2001.
XX
PF   29-DEC-2000; 2000WO-IT000554.
XX
PR   30-DEC-1999;   99IT-RM000803.
XX
PA   (CNEN ) ENEA ENTE NUOVE TECNOLOGIE ENERGIA.
PA   (CONS-) SOC CONSORTILE METAPONTUM AGROBIOS SRL.
XX
PI   Benvenuto E,  Franconi R,  Desiderio A,  Tavladoraki P;
XX
DR   WPI; 2001-502555/55.
DR   N-PSDB; AAS11887.
XX
PT   Peptides which are able to confer stability and solubility to an antibody
PT   comprising these peptides, useful for treating pathologies (e.g. tumor)
PT   associated with accumulation of a molecule inside or outside a human, or
PT   animal cell.
XX
PS   Example 1; Page 81; 109pp; English.
XX
CC   The invention relates to peptides which are able to confer stability and
CC   solubility to an antibody comprising these peptides. The peptides are
CC   especially H-FR1, H-FR2, HF-R3, HFR4, L-FR1, L-FR2, L-FR3 or L-FR4
CC   present within a variable region of an antibody which makes the antibody
CC   soluble and stable in cytoplasm. Peptides having the sequences of HFR1 to
CC   H-FR4 are present within the variable region of the heavy chain of an
CC   antibody, covalently linked to the H-CDR1, H-CDR2, H-CDR3 in the order (H
CC   -FR1)-(H-CDR1)-(H-FR2)-(H-CDR2)-(H-FR3)-(H-CDR3)-(H-CDR4) and peptides
CC   having the sequences of L-FR1 to L-FR4 are present within the variable
CC   region of the light chain of an antibody, covalently linked to the L-
```

```
CC   CDR1, L-CDR2, L-CDR3 in the order (L-FR1)-( L-CDR1)-( L-FR2)-( L-CDR2)-(
CC   L-FR3)-( L-CDR3)-( L-CDR4). The antibodies and polynucleotides (e.g. by
CC   gene therapy) are useful for the manufacture of a medicament for the
CC   treatment of pathologies associated with accumulation of a molecule
CC   inside or outside a human, animal cell or plant cell. The pathologies are
CC   infectious (e.g. viral infections such as HIV, human immunodeficiency
CC   virus, infections), tumour, metabolic and immune (especially auto-immune)
CC   pathologies. The present sequence represents the synthetic antibody
CC   scFv(F8) which is used as a basis for constructing synthetic antibodies
CC   incorporating the peptides of the invention
XX
SQ   Sequence 252 AA;

  Query Match             72.8%;  Score 926.5;  DB 4;  Length 252;
  Best Local Similarity   72.6%;  Pred. No. 1.4e-58;
  Matches  180;  Conservative   26;  Mismatches   31;  Indels   11;  Gaps    3;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              ||:|||||| ||:||||||||||||||||||||||||||||||| ||||:| |  |:|
Db          1 QVQLQESGGDLVQPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLELVATINSNGGSTFY 60

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCA-RGNWEGW-----YFDVWGQGTT 114
              |||||||||||||||||||||||||||||||||||| | |:  :     ||| ||||||
Db         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARRRNYPYYYGSRGYFDYWGQGTT 120

Qy        115 VTVSSGGGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSI-----RYIYWY 169
              ||||||||||||||||||:|||||||| :: | |:| |::| || |:      :::||
Db        121 VTVSSGGGGSGGGGSGGGGSDIELTQSPASLAVSLGQRATISCRASESVDSYGNSFMHWY 180

Qy        170 QQKPGSSPRLLIYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPY 229
              |||||  |:||||   |:  |:| ||||||| | ::|||| :||:| ||||||: :  |:
Db        181 QQKPGQPPKLLIYRALNLESGIPARFSGSGSRTDFTLTINPVEADDVATYYCQQSNEDPW 240

Qy        230 TFGGGTKL 237
              ||||||||
Db        241 TFGGGTKL 248


Search completed: December  8, 2004, 17:16:45
Job time : 169.707 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2004, 17:06:18 ; Search time 46.2439 Seconds
(without alignments)
493.111 Million cell updates/sec

Title: US-10-073-301A-9
Perfect score: 1272
Sequence: 1 QVKLQESGGGLVKPGGSLKL.....TYYCQEWSGYPYTFGGGTKL 237

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	792.5	62.3	268	2	A56446	Ig heavy chain V r
2	697	54.8	249	2	S41374	single chain Fv an
3	631	49.6	233	2	JC5322	p53 specific singl
4	542	42.6	120	2	S55536	Ig heavy chain V r
5	535.5	42.1	119	2	F27888	Ig heavy chain V r
6	533	41.9	120	2	S55538	Ig heavy chain V r
7	533	41.9	120	2	S55537	Ig heavy chain V r
8	532	41.8	120	2	S55539	Ig heavy chain V r
9	532	41.8	548	2	S38864	Ig epsilon chain C
10	530	41.7	122	2	E27888	Ig heavy chain V r
11	522	41.0	254	2	B31790	Ig heavy chain V r
12	520.5	40.9	112	2	S26327	Ig heavy chain V r
13	520	40.9	117	2	PL0249	Ig heavy chain V r
14	514.5	40.4	121	2	I27887	Ig heavy chain V r
15	513	40.3	117	2	PL0252	Ig heavy chain V r
16	510.5	40.1	118	2	PH0096	Ig heavy chain V r
17	510.5	40.1	119	2	B27889	Ig heavy chain V r
18	510	40.1	152	2	B26471	Ig heavy chain pre
19	509.5	40.1	119	2	D27889	Ig heavy chain V r
20	509	40.0	124	2	C27888	Ig heavy chain V r
21	506.5	39.8	118	2	PH0097	Ig heavy chain V r
22	505.5	39.7	121	2	S55540	Ig heavy chain V r
23	505	39.7	138	2	S09258	Ig heavy chain V r
24	503.5	39.6	121	2	H27888	Ig heavy chain V r
25	501.5	39.4	123	2	G27888	Ig heavy chain V r
26	497.5	39.1	139	2	S38808	Ig heavy chain - m
27	495	38.9	119	2	B34353	anti-peptide Fab'
28	493.5	38.8	121	2	A27888	Ig heavy chain V r
29	493.5	38.8	121	2	D27888	Ig heavy chain V r
30	493	38.8	108	2	PH1006	Ig heavy chain V r
31	493	38.8	118	2	S20641	Ig heavy chain V r
32	492.5	38.7	112	2	A27889	Ig heavy chain V r
33	491	38.6	119	2	PH0098	Ig heavy chain V r
34	488.5	38.4	123	2	S63597	Ig heavy chain, V
35	483.5	38.0	120	2	S12953	Ig heavy chain V r
36	482.5	37.9	108	2	PH1010	Ig heavy chain V r
37	482.5	37.9	121	2	B27888	Ig heavy chain V r
38	482	37.9	138	2	S31666	Ig heavy chain V r
39	481	37.8	119	2	A43413	Ig heavy chain V r
40	480.5	37.8	121	2	H27887	Ig heavy chain V r
41	480.5	37.8	141	2	S31669	Ig heavy chain V r
42	479	37.7	119	2	D36005	Ig heavy chain V r
43	478	37.6	142	2	C34903	Ig heavy chain pre
44	477	37.5	111	2	PH1007	Ig heavy chain V r
45	476.5	37.5	147	2	I37780	Ig variable region

ALIGNMENTS

RESULT 1
A56446
Ig heavy chain V region (3H-3H scFv) - mouse (strain BALB/C)
C;Species: Mus musculus (house mouse)
C;Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 16-Aug-1996
C;Accession: A56446
R;Tang, P.M.; Foltz, L.A.; Mahoney, W.C.; Schueler, P.A.
J. Biol. Chem. 270, 7829-7835, 1995
A;Title: A high affinity digoxin-binding protein displayed on M13 is functionally ident
A;Reference number: A56446; MUID:95229583; PMID:7713873
A;Accession: A56446
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-268 <TAN>
A;Cross-references: GB:U20617
C;Keywords: heterotetramer; immunoglobulin

Query Match 62.3%; Score 792.5; DB 2; Length 268;
Best:Local Similarity 64.1%; Pred. No. 8.8e-48;
Matches 152; Conservative 29; Mismatches 55; Indels 1; Gaps 1;

```
Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              |||||||  ||||| |:||||  |||       | ||:| |:: |||:  |:     | |
Db          3 QVKLQESGAELVKPGASVKLSCTTSGFNIKDTYMHWVKQRPEQGLEWIGRIAPANGITKY 62

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
                :|: ||: | : || |||:||| |||||:||||   :   | : ||||||||||||
Db         63 DPKFQGKATIAADTSSNTAYLQLSSLTSEDTAVYYCA-SYYLTRYENYWGQGTTVTVSSG 121

Qy        121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRLL 180
              ||||||| ||||||:|||||||||||||| ||:|||:| ||||: :|||||||  :||:|
Db        122 GGGSGGGDSGGGGSDIELTQSPAIMSASLGEKVTMSCRASSSVNFIYWYQQKSDASPKLW 181

Qy        181 IYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
              :| ||:: |||| |||||||| ||||||: || |||||||||:::  |:||| ||||
Db        182 VYYTSHLPPGVPARFSGSGSGNSYSLTISSMEGEDAATYYCQQFTSSPFTFGSGTKL 238
```

RESULT 2
S41374
single chain Fv antibody - mouse
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995
C;Accession: S41374
R;Artsaenko, O.; Weiler, E.W.; Muentz, K.; Conrad, U.
submitted to the EMBL Data Library, January 1994
A;Description: Construction and functional characterization of a single chain Fv antibo
A;Reference number: S41374
A;Accession: S41374
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-249 <ART>

```
A;Cross-references: EMBL:Z29480

  Query Match             54.8%;  Score 697;  DB 2;  Length 249;
  Best Local Similarity   55.6%;  Pred. No. 3.1e-41;
  Matches  135;  Conservative   35;  Mismatches   67;  Indels    6;  Gaps    1;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              ||:||:||  ||:|| |:|||| |||| |    : ||:| |:| |||:| |:       |
Db          1 QVQLQQSGAELVRPGASVKLSCTASGFNFKDDYIHWVKQRPEKGLEWIARIAPASGNVKY 60

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
                 : : ||: | : || || :||| |||||:||||| :        ||||:||||||
Db         61 VPRFQDKATITADTSSNTAYLLLSSLTSEDTAVYYCARRDTLYTSLGYWGQGSTVTVSSR 120

Qy        121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSI------RYIYWYQQKPG 174
              ||||||||||||||:|||||||  :   ||| |:::| :| |:       |::|: |:||
Db        121 GGGSGGGGSGGGGSDIELTQSPPSVVVIPGESVSISCRSSKSLLYSDGDSYLFWFLQRPG 180

Qy        175 SSPRLLIYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGG 234
              ||:||||  ||:| ||| |||||||||::| |:|:||||   ||| :   || ||| |
Db        181 QSPQLLIYRMSNLASGVPDRFSGSGSGTSFTLRISRVEAEDVGVYYCMQHREYPLTFGAG 240

Qy        235 TKL 237
              |||
Db        241 TKL 243


RESULT 3
JC5322
p53 specific single-chain antibody Pab421 - human
C;Species: Homo sapiens (man)
C;Date: 15-May-1997 #sequence_revision 15-May-1997 #text_change 18-Jul-1997
C;Accession: JC5322
R;Jannot, C.B.; Hynes, N.E.
Biochem. Biophys. Res. Commun. 230, 242-246, 1997
A;Title: Characterization of scFv-421, a single-chain antibody targeted to p53.
A;Reference number: JC5322; MUID:97168950; PMID:9016757
A;Accession: JC5322
A;Molecule type: mRNA
A;Residues: 1-233 <JAN>
A;Experimental source: hydricloma cell
C;Comment: This protein specifically binds the tumor suppressor protein p53. It restores

  Query Match             49.6%;  Score 631;  DB 2;  Length 233;
  Best Local Similarity   54.6%;  Pred. No. 1e-36;
  Matches  131;  Conservative   30;  Mismatches   63;  Indels   16;  Gaps    5;

Qy          5 QESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATI--SSGGSYTYYPD 62
              ||||  ||: | |:|||  |||  : | | ||:: |:: |||:  |   :| :
Db          1 QESGAELVRSGASVKLSCTTSGFNINDYYMHWVKKRPEQGLEWIGRIDPENGDADMTRSS 60

Qy         63 SVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSGGG 122
                ||   |:: | : || ||:|||| |||||:|||  |       | ||||||||||||
Db         61 GVKA--TMTADTSSNTAYLQLSSLTSEDTAVYYCNAG------MDYWGQGTTVTVSSGGG 112

Qy        123 GSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSI-----RYIYWYQQKPGSSP 177
              ||||  ||||||:|||||||| :: | |:| |::| || |:      |::| |||||  |
Db        113 GSGGRASGGGGSDIELTQSPASLAVSLGQRATISCRASKSVSTSGYSYMHWNQQKPGQPP 172

Qy        178 RLLIYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
              |||||  ||:  ||| ||||||||| ::| |: :| |||||||||       :  |||||
Db        173 RLLIYLVSNLESGVPARFSGSGSGTDFTLNIHPVEEEDAATYYCQHIRELTRS-EGGTKL 231


RESULT 4
S55536
Ig heavy chain V region pe20 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 23-Jul-1999
C;Accession: S55536
R;Boettger, V.; Boettger, A.; Lane, E.B.; Spruce, B.A.
J. Mol. Biol. 247, 932-946, 1995
A;Title: Comprehensive epitope analysis of monoclonal anti-proenkephalin antibodies usin
utations in the variable region genes.
A;Reference number: S55528; MUID:95239763; PMID:7536850
A;Accession: S55536
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-120 <BOE>
A;Cross-references: EMBL:X82589; NID:g854304; PIDN:CAA57925.1; PID:g854305
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;14-97/Domain: immunoglobulin homology <IMM>

  Query Match             42.6%;  Score 542;  DB 2;  Length 120;
  Best Local Similarity   86.7%;  Pred. No. 7.1e-31;
  Matches  104;  Conservative    8;  Mismatches    6;  Indels    2;  Gaps    1;

Qy          2 VKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYYP 61
              |:|||||||||||||||||||||||||||:| |||||||:|||||||||||||||||||
Db          1 VQLQESGGGLVKPGGSLKLSCAASGFTFSTYAMSWVRQTPEKRLEWVATISSGGSYTYYP 60

Qy         62 DSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCAR--GNWEGWYFDVWGQGTTVTVSS 119
              |||||||||||||||||||||||||:|||||||||||   ::: :  | ||||||||||
Db         61 DSVKGRFTISRDNAKNTLYLQMSSLRSEDTAMYYCARLYYDYDPYVMDYWGQGTTVTVSS 120


RESULT 5
F27888
Ig heavy chain V region (H158-89H4) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 16-Aug-1996
C;Accession: F27888
R;Caton, A.J.; Brownlee, G.G.; Staudt, L.M.; Gerhard, W.
EMBO J. 5, 1577-1587, 1986
A;Title: Structural and functional implications of a restricted antibody response to a
A;Reference number: A91043; MUID:86300658; PMID:2427335
A;Accession: F27888
A;Molecule type: DNA
A;Residues: 1-119 <CAT>
A;Experimental source: strain Balb/c
A;Note: this sequence was determined from the germline gene
C;Comment: This chain was isolated from a hybridoma protein that binds influenza virus
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

  Query Match             42.1%;  Score 535.5;  DB 2;  Length 119;
  Best Local Similarity   88.3%;  Pred. No. 2e-30;
  Matches  106;  Conservative    3;  Mismatches    6;  Indels    5;  Gaps    2;

Qy          2 VKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYYP 61
              ||| |||||||||||||||||||||||||||||| ||||||||:|||||||||||||||||
Db          2 VKLVESGGGLVKPGGSLKLSCAASGFTFSSYTMSWVRQTPEKRLEWVATISSGGSYTYYP 61

Qy         62 DSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWY---FDVWGQGTTVTVS 118
              |||||||||||||||||||||||||||||||||||| ||  | :|    | |||||:||||
Db         62 DSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCPRG--ETYYDYAMDYWGQGTSVTVS 119


RESULT 6
S55538
Ig heavy chain V region pe22 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 23-Jul-1999
C;Accession: S55538
R;Boettger, V.; Boettger, A.; Lane, E.B.; Spruce, B.A.
J. Mol. Biol. 247, 932-946, 1995
A;Title: Comprehensive epitope analysis of monoclonal anti-proenkephalin antibodies usi
utations in the variable region genes.
A;Reference number: S55528; MUID:95239763; PMID:7536850
A;Accession: S55538
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-120 <BOE>
A;Cross-references: EMBL:X82591; NID:g854308; PIDN:CAA57927.1; PID:g854309
```

```
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;14-97/Domain: immunoglobulin homology <IMM>

  Query Match             41.9%;  Score 533;  DB 2;  Length 120;
  Best Local Similarity   85.0%;  Pred. No. 3e-30;
  Matches  102;  Conservative    8;  Mismatches    8;  Indels    2;  Gaps    1;

Qy          2 VKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYYP 61
              |||||||||||||||||||||||||||||| ||||||||:|||||||||||| |:|||||
Db          1 VKLQESGGGLVKPGGSLKLSCAASGFTFSSYAMSWVRQTPEKRLEWVATISSSGNYTYYP 60

Qy         62 DSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCAR--GNWEGWYFDVWGQGTTVTVSS 119
              ||||||||:|||||||||||||||:|||||||| |   ::: :  | ||||||||||
Db         61 DSVKGRFTMSRDNAKNTLYLQMSSLRSEDTAMYYCVRLYYDYDPYVMDFWGQGTTVTVSS 120


RESULT 7
S55537
Ig heavy chain V region pe21 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 23-Jul-1999
C;Accession: S55537
R;Boettger, V.; Boettger, A.; Lane, E.B.; Spruce, B.A.
J. Mol. Biol. 247, 932-946, 1995
A;Title: Comprehensive epitope analysis of monoclonal anti-proenkephalin antibodies usin
utations in the variable region genes.
A;Reference number: S55528; MUID:95239763; PMID:7536850
A;Accession: S55537
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-120 <BOE>
A;Cross-references: EMBL:X82590; NID:g854306; PIDN:CAA57926.1; PID:g854307
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;14-97/Domain: immunoglobulin homology <IMM>

  Query Match             41.9%;  Score 533;  DB 2;  Length 120;
  Best Local Similarity   85.8%;  Pred. No. 3e-30;
  Matches  103;  Conservative    7;  Mismatches    8;  Indels    2;  Gaps    1;

Qy          2 VKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYYP 61
              |:|||||||||||||||||||| |||||| ||||||||:||||||||||||||||||||
Db          1 VQLQESGGGLVKPGGSLKLSCAASRFTFSSYAMSWVRQTPEKRLEWVATISSGGSYTYYP 60

Qy         62 DSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCAR--GNWEGWYFDVWGQGTTVTVSS 119
              ||||||||||||||||||||||:|||||||| |   ::: :  | ||||||||||
Db         61 DSVKGRFTISRDNAKNTLYLQMSSLRSEDTAMYYCTRLYYDYDPYVMDYWGQGTTVTVSS 120


RESULT 8
S55539
Ig heavy chain V region pe24 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 23-Jul-1999
C;Accession: S55539
R;Boettger, V.; Boettger, A.; Lane, E.B.; Spruce, B.A.
J. Mol. Biol. 247, 932-946, 1995
A;Title: Comprehensive epitope analysis of monoclonal anti-proenkephalin antibodies usin
utations in the variable region genes.
A;Reference number: S55528; MUID:95239763; PMID:7536850
A;Accession: S55539
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-120 <BOE>
A;Cross-references: EMBL:X82593; NID:g854312; PIDN:CAA57929.1; PID:g854313
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;14-97/Domain: immunoglobulin homology <IMM>

  Query Match             41.8%;  Score 532;  DB 2;  Length 120;
  Best Local Similarity   85.0%;  Pred. No. 3.5e-30;
  Matches  102;  Conservative    8;  Mismatches    8;  Indels    2;  Gaps    1;

Qy          2 VKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYYP 61
              |:|||||||||||||||||||||||||||| ||||||||:|||||||||||| |||||.
Db          1 VQLQESGGGLVKPGGSLKLSCAASGFTFSSYAMSWVRQTPEKRLEWVATISSSGSYTYYP 60

Qy         62 DSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCAR--GNWEGWYFDVWGQGTTVTVSS 119
              ||||||||:|||||||||||||||:|||||||| |   ::: :  | ||||||||||
Db         61 DSVKGRFTMSRDNAKNTLYLQMSSLRSEDTAMYYCVRLYYDYDPYVMDYWGQGTTVTVSS 120


RESULT 9
S38864
Ig epsilon chain C region - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-May-2001
C;Accession: S38864
R;Kipp, B.; Becker, W.; Schlaak, M.
submitted to the EMBL Data Library, November 1993
A;Description: Combination of a defined specificity and desired isotype by cloning of a
A;Reference number: S38864
A;Accession: S38864
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-548 <KIP>
A;Cross-references: EMBL:Z27397; NID:g416537; PIDN:CAA81788.1; PID:g940782
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;353-421/Domain: immunoglobulin homology <IMM>

  Query Match             41.8%;  Score 532;  DB 2;  Length 548;
  Best Local Similarity   68.3%;  Pred. No. 1.7e-29;
  Matches  112;  Conservative    6;  Mismatches   22;  Indels   24;  Gaps    3;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              |||| |||| |||||||||||||||| |||||||||||| ||||||||||||||:||||
Db          1 QVKLLESGGDLVKPGGSLKLSCAASGLTFSSYGMSWVRQIPDKRLEWVATISSGGTYTYY 60

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCAR-GNWEGWYFDVWGQGTTVTVSS 119
              |||||||||||||||||||||||||||||||||||||| |       |  ||||| ||||:
Db         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARQGVSTMIRFAYWGQGTLVTVSA 120

Qy        120 GGGGSGGGGSGGGGSNIELTQSPAIMSASPG------ERVTMTC 157
              |                    |  |::   :||         ||: |
Db        121 G----------------KTTPPSVYPLAPGSAAQTNSMVTLGC 147


RESULT 10
E27888
Ig heavy chain V region (H35-C6) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 16-Aug-1996
C;Accession: E27888
R;Caton, A.J.; Brownlee, G.G.; Staudt, L.M.; Gerhard, W.
EMBO J. 5, 1577-1587, 1986
A;Title: Structural and functional implications of a restricted antibody response to a
A;Reference number: A91043; MUID:86300658; PMID:2427335
A;Accession: E27888
A;Molecule type: DNA
A;Residues: 1-122 <CAT>
A;Experimental source: strain Balb/c
A;Note: this sequence was determined from the germline gene
C;Comment: This chain was isolated from a hybridoma protein that binds influenza virus
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

  Query Match             41.7%;  Score 530;  DB 2;  Length 122;
  Best Local Similarity   85.1%;  Pred. No. 4.9e-30;
  Matches  103;  Conservative    5;  Mismatches    9;  Indels    4;  Gaps    1;

Qy          2 VKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYYP 61
              ||| |||||||||||||||||||||||||| ||||||||:|||||||||||||||||||
Db          2 VKLVESGGGLVKPGGSLKLSCAASGFTFSSYTMSWVRQTPEKRLEWVATISSGGSYTYYP 61
```

```
Qy          62 DSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARG----NWEGWYFDVWGQGTTVTV 117
               ||||||||||||||||||||||||||||||||||| |  .    :: :  | |||||:|||
Db          62 DSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCTRQGEGYRYDDYAMDYWGQGTSVTV 121

Qy         118 S 118
               |
Db         122 S 122
```

RESULT 11
B31790
Ig heavy chain V region (17/9) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 23-May-1997
C;Accession: B31790
R;Schulze-Gahmen, U.; Rini, J.M.; Arevalo, J.; Stura, E.A.; Kenten, J.H.; Wilson, I.A.
J. Biol. Chem. 263, 17100-17105, 1988
A;Title: Preliminary crystallographic data, primary sequence, and binding data for an an
A;Reference number: A92686; MUID:89034213; PMID:3182835
A;Accession: B31790
A;Molecule type: mRNA
A;Residues: 1-254 <SCH>
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

  Query Match   41.0%; Score 522; DB 2; Length 254;
  Best Local Similarity   64.5%; Pred. No. 3.7e-29;
  Matches 107; Conservative   13; Mismatches   20; Indels   26; Gaps   3;

```
Qy           1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
               :|:| |||| |||||||||||||||:|||||||||||||||||||||||:|| ||||
Db           1 EVQLVESGGDLVKPGGSLKLSCAASGFSFSSYGMSWVRQTPDKRLEWVATISNGGGYTYY 60

Qy          61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCA-RGNWEGWYFDVWGQGTTVTVSS 119
               |||||||||||||||||||||||||||||:|||||| |  ::   |  ||||| ||||:
Db          61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDSAMYYCARRERYDENGFAYWGQGTLVTVSA 120

Qy         120 GGGGSGGGGSGGGGSNIELTQSPAIMSASP--------GERVTMTC 157
                           | :|::   :|          |  ||: |
Db         121 A-----------------KTTAPSVYPLAPVCGXXDTTGSSVTLGC 149
```

RESULT 12
S26327
Ig heavy chain V region - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1998 #sequence_revision 19-Mar-1998 #text_change 21-Jan-2000
C;Accession: S26327
R;Stark, S.E.; Caton, A.J.
J. Exp. Med. 174, 613-624, 1991
A;Title: Antibodies that are specific for a single amino acid interchange in a protein e
A;Reference number: S26309; MUID:91341421; PMID:1908510
A;Accession: S26327
A;Molecule type: mRNA
A;Residues: 1-112 <STA>
A;Cross-references: EMBL:X59192
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;9-91/Domain: immunoglobulin homology <IMM>

  Query Match   40.9%; Score 520.5; DB 2; Length 112;
  Best Local Similarity   88.7%; Pred. No. 2e-29;
  Matches 102; Conservative   5; Mismatches   3; Indels   5; Gaps   3;

```
Qy           7 SGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYYPDSVKG 66
               |||||||||||||||||||||||| ||||||||:|||||||:|||||| |||||||||
Db           1 SGGGLVKPGGSLKLSCAASGFTFSSYAMSWVRQTPEKRLEWVASISSGGS-TYYPDSVKG 59

Qy          67 RFTISRDNAKNTLYLQMSSLKSEDTAMYYCAR--GNWEGWYFDVWGQGTTVTVSS 119
               |||||||||:| ||||||||:||||||||||  ||:  |||||||| ||||||||
Db          60 RFTISRDNARNILYLQMSSLRSEDTAMYYCARLYGNY--WYFDVWGAGTTVTVSS 112
```

RESULT 13
PL0249
Ig heavy chain V region (anti-DNA, 3E12VH) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 16-Aug-1996
C;Accession: PL0249
R;Shlomchik, M.; Mascelli, M.; Shan, H.; Radic, M.Z.; Pisetsky, D.; Marshak-Rothstein, .
J. Exp. Med. 171, 265-297, 1990
A;Title: Anti-DNA antibodies from autoimmune mice arise by clonal expansion and somatic
A;Reference number: PL0231; MUID:90111618; PMID:2104919
A;Accession: PL0249
A;Molecule type: mRNA
A;Residues: 1-117 <SHL>
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;1-30/Region: framework 1
F;15-98/Domain: immunoglobulin homology <IMM>
F;31-35/Region: complementarity-determining 1
F;36-49/Region: framework 2
F;50-66/Region: complementarity-determining 2
F;67-98/Region: framework 3
F;99-108/Region: complementarity-determining 3
F;109-117/Region: framework 4

  Query Match   40.9%; Score 520; DB 2; Length 117;
  Best Local Similarity   83.8%; Pred. No. 2.3e-29;
  Matches   98; Conservative   7; Mismatches   12; Indels   0; Gaps   0;

```
Qy           1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
               :||| |||||||||||||||||||||||||| |||||||| ||||||| ||  |  |||
Db           1 EVKLVESGGGLVKPGGSLKLSCAASGFTFSSYTMSWVRQTPAKRLEWVANISRRGGSTYY 60

Qy          61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTV 117
               |||||||||||||||:||||||||||:|||||:||||| ::  |:||||| ||||||
Db          61 PDSVKGRFTISRDNARNTLYLQMSSLRSEDTAVYYCARRDYSHWFFDVWGAGTTVTV 117
```

RESULT 14
I27887
Ig heavy chain V region (H37-45) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 16-Aug-1996
C;Accession: I27887
R;Caton, A.J.; Brownlee, G.G.; Staudt, L.M.; Gerhard, W.
EMBO J. 5, 1577-1587, 1986
A;Title: Structural and functional implications of a restricted antibody response to a
A;Reference number: A91043; MUID:86300658; PMID:2427335
A;Accession: I27887
A;Molecule type: DNA
A;Residues: 1-121 <CAT>
A;Experimental source: strain Balb/c
A;Note: this sequence was determined from the germline gene
C;Comment: This chain was isolated from a hybridoma protein that binds influenza virus
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

  Query Match   40.4%; Score 514.5; DB 2; Length 121;
  Best Local Similarity   82.6%; Pred. No. 5.7e-29;
  Matches 100; Conservative   6; Mismatches   12; Indels   3; Gaps   1;

```
Qy           1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
               :| | |||||||||||||||||||||| | ||||||||:|||||||||||||:||||
Db           1 EVMLVESGGGLVKPGGSLKLSCAASGFTFSIYAMSWVRQTPEKRLEWVATISSGGTYTYY 60

Qy          61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARG---NWEGWYFDVWGQGTTVTV 117
               |||||||||||||||||||||||||||||||||||||      | :  | |||||:|||
Db          61 PDSVKGRFTISRDNAKNTLYLQMSRLRSEDTAMYYCAREEGLRLEDYAMDYWGQGTSVTV 120

Qy         118 S 118
               |
Db         121 S 121
```

```
RESULT 15
PL0252
Ig heavy chain V region (anti-DNA, clones 2E3VH, 6B8VH, and 3G9VH) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 16-Aug-1996
C;Accession: PL0252; PL0251
R;Shlomchik, M.; Mascelli, M.; Shan, H.; Radic, M.Z.; Pisetsky, D.; Marshak-Rothstein, A
J. Exp. Med. 171, 265-297, 1990
A;Title: Anti-DNA antibodies from autoimmune mice arise by clonal expansion and somatic
A;Reference number: PL0231; MUID:90111618; PMID:2104919
A;Accession: PL0252
A;Molecule type: mRNA
A;Residues: 1-117 <SHL>
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;1-30/Region: framework 1
F;15-98/Domain: immunoglobulin homology <IMM>
F;31-35/Region: complementarity-determining 1
F;36-49/Region: framework 2
F;50-66/Region: complementarity-determining 2
F;67-98/Region: framework 3
F;99-108/Region: complementarity-determining 3
F;109-117/Region: framework 4

  Query Match             40.3%;  Score 513;  DB 2;  Length 117;
  Best Local Similarity   82.9%;  Pred. No. 6.9e-29;
  Matches   97;  Conservative    7;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              :||| |||||||||||||||||||||||| |||| ||| ||||||| ||  |  |||
Db          1 EVKLVESGGGLVKPGGSLKLSCAASGFTFSSYTMSWVGQTPAKRLEWVANISRRGGSTYY 60

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTV 117
              |||||||||||||||:||||||||||:|||||:|||||  ::  |:|||||  ||||||
Db         61 PDSVKGRFTISRDNARNTLYLQMSSLRSEDTAVYYCARRDYSHWFFDVWGAGTTVTV 117


Search completed: December  8, 2004, 17:13:38
Job time : 47.2439 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2004, 17:06:18 ; Search time 6.29268 Seconds
(without alignments)
513.066 Million cell updates/sec

Title: US-10-073-301A-2
Perfect score: 45
Sequence: 1 IMDQVPFSV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	100.0	9	2	AAR84817	Aar84817 Modified
2	45	100.0	9	4	AAU28966	Aau28966 Modified
3	45	100.0	9	4	AAU72002	Aau72002 gp100 mel
4	45	100.0	9	4	AAU10221	Aau10221 Cancer ce
5	45	100.0	9	4	AAE05120	Aae05120 Modified
6	45	100.0	9	5	AAO17085	Aao17085 Gp 100 an
7	45	100.0	9	5	AAE17297	Aae17297 gp100-mod
8	45	100.0	9	6	AAE36055	Aae36055 Tumour as
9	45	100.0	9	6	AAG79857	Aag79857 HLA-A2-re
10	45	100.0	9	6	AAE35575	Aae35575 Melanoma
11	45	100.0	9	6	ABJ19877	Abj19877 MHC bindi
12	45	100.0	9	6	ABU08668	Abu08668 Cancer ce
13	45	100.0	9	6	ADA89157	Ada89157 Gp100-der
14	45	100.0	9	7	AAE38651	Aae38651 HLA-A2 re
15	45	100.0	9	7	ADB97733	Adb97733 Human gp1
16	45	100.0	17	4	AAU72227	Aau72227 gp100-der
17	45	100.0	17	4	AAU72226	Aau72226 gp100-der
18	45	100.0	17	4	AAU72231	Aau72231 gp100-der
19	45	100.0	20	4	AAU72223	Aau72223 gp100-der
20	45	100.0	20	4	AAU72224	Aau72224 gp100-der
21	45	100.0	20	4	AAU72229	Aau72229 gp100-der
22	45	100.0	20	5	AAE13455	Aae13455 Human gp1
23	45	100.0	20	5	AAE13456	Aae13456 Human gp1
24	45	100.0	25	4	AAU72228	Aau72228 gp100-der
25	45	100.0	25	4	AAU72232	Aau72232 gp100-der
26	45	100.0	30	4	AAB61648	Aab61648 Gp100 pep
27	45	100.0	31	4	AAU72230	Aau72230 gp100-der
28	45	100.0	31	4	AAU72225	Aau72225 gp100-der
29	45	100.0	31	5	AAE13457	Aae13457 Human gp1
30	45	100.0	35	6	AAE35578	Aae35578 TA peptid
31	45	100.0	661	4	AAB97816	Aab97816 Modified
32	45	100.0	661	4	AAE05116	Aae05116 Modified
33	45	100.0	661	4	AAB98206	Aab98206 Human gp1
34	42	93.3	9	2	AAR84816	Aar84816 Modified
35	42	93.3	9	4	AAU28965	Aau28965 Modified
36	42	93.3	9	7	ADB97732	Adb97732 Human gp1
37	41	91.1	9	2	AAR84818	Aar84818 Modified
38	41	91.1	9	4	AAU28967	Aau28967 Modified
39	41	91.1	9	4	AAB97708	Aab97708 Avipox vi
40	41	91.1	9	4	AAB98098	Aab98098 Modified
41	41	91.1	9	7	ADB97734	Adb97734 Human gp1
42	39	86.7	9	2	AAR84210	Aar84210 gp100 mel
43	39	86.7	9	2	AAW70018	Aaw70018 Melanoma-
44	39	86.7	9	2	AAW54599	Aaw54599 Peptide 4
45	39	86.7	9	2	AAW78851	Aaw78851 PMEL 17 (

ALIGNMENTS

RESULT 1
AAR84817
ID AAR84817 standard; peptide; 9 AA.
XX
AC AAR84817;
XX
DT 25-APR-1996 (first entry)
XX
DE Modified gp100 melanocyte-melanoma specific antigenic peptide G9-209-2M.
XX
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;
KW metastatic melanoma; tumour-associated antigen; immunogenic peptide;
KW diagnosis; prognosis; prophylaxis; therapy; vaccine.
XX
OS Synthetic.
XX
PN WO9529193-A2.
XX
PD 02-NOV-1995.
XX
PF 21-APR-1995; 95WO-US005063.
XX
PR 22-APR-1994; 94US-00231565.
PR 05-APR-1995; 95US-00417174.
XX
PA (USSH) US SEC DEPT HEALTH.
XX
PI Kawakami Y, Rosenberg SA;
XX
DR WPI; 1995-382963/49.
XX
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and immunise
PT animal against melanoma.
XX
PS Example 5; Page 107; 184pp; English.
XX
CC AAR84816-836 are G9-209 peptides modified to improve immunogenicity. G9-
CC 209 is an immunogenic peptide based on the melanoma derived antigen,
CC gp100 (see AAR84210). The peptides are used in medicaments for the
CC treatment or prevention (by immunization) of melanoma. Antibodies against
CC MART-1 and its immunogenic peptides may be used in the detection and
CC isolation of MART-1 from a sample, the detection of which is indicative
CC of a disease state (melanoma or metastatic melanoma)
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 45; DB 2; Length 9;

```
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              |||||||||
Db          1 IMDQVPFSV 9


RESULT 2
AAU28966
ID   AAU28966 standard; peptide; 9 AA.
XX
AC   AAU28966;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Modified gp100 G9-209 peptide #2.
XX
KW   Human; MART-I; immunogenic; melanoma antigen recognised by T lymphocyte;
KW   diagnostic; therapeutic; vaccine; melanoma; in vivo tumour recognition;
KW   in vivo tumour rejection.
XX
OS   Synthetic.
XX
PN   US6270778-B1.
XX
PD   07-AUG-2001.
XX
PF   12-MAR-1999;   99US-00267439.
XX
PR   22-APR-1994;   94US-00231565.
PR   05-APR-1995;   95US-00417174.
PR   05-MAY-1998;   98US-00073138.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Kawakami Y,  Rosenberg SA;
XX
DR   WPI; 2001-595403/67.
XX
PT   Immunogenic peptide useful in vaccines comprises specific amino acids of
PT   new melanoma antigen recognized by T lymphocytes.
XX
PS   Example 5; Col 55; 73pp; English.
XX
CC   The invention relates to a novel immunogenic peptide comprising 5-20
CC   contiguous amino acids of new melanoma antigen recognised by T
CC   lymphocytes (MART-1). The peptide sequence contains at least one amino
CC   acid modification of MART-1. The peptide is used in diagnostic and
CC   therapeutic methods as an immunogen or vaccine to prevent or treat
CC   melanoma, and for in vivo tumour recognition and rejection. AAU28888-
CC   AAU29008 represent MART-1 peptide amino acid sequences, and related
CC   sequences of the invention
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 45;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              |||||||||
Db          1 IMDQVPFSV 9


RESULT 3
AAU72002
ID   AAU72002 standard; peptide; 9 AA.
XX
AC   AAU72002;
XX
DT   26-FEB-2002  (first entry)
XX
DE   gp100 melanoma antigen #14.
XX
KW   Melanoma antigen; MART-1; MAGE-1; gp100; cytostatic; immune response;
KW   immunotherapeutic; heat shock protein; tyrosinase; BAGE; NYEs01; GM2;
KW   tyrosinase related protein 1; tyrosinase related protein 2; vaccine;
KW   javelin molecule; melanoma antigen recognised by T cells-1; human.
XX
OS   Homo sapiens.
XX
PN   WO200178655-A2.
XX
PD   25-OCT-2001.
XX
PF   17-APR-2001; 2001WO-US012449.
XX
PR   17-APR-2000; 2000US-0197462P.
XX
PA   (HOUG/) HOUGHTON A.
PA   (LIVI/) LIVINGSTON P.
PA   (ALAW/) AL-AWQATI Q.
PA   (MAYH/) MAYHEW M.
PA   (HOEM/) HOE M.
XX
PI   Houghton A,  Livingston P,  Al-Awqati Q,  Mayhew M,  Hoe M;
XX
DR   WPI; 2001-663092/76.
XX
PT   Anti cancer vaccine for the treatment of melanoma comprises a heat shock
PT   protein and a melanoma antigen i.e. tyrosinase.
XX
PS   Claim 2; Page 11; 150pp; English.
XX
CC   The invention relates to a method of induction of an immune response,
CC   comprising administration of an immunotherapeutic composition, comprising
CC   a heat shock protein, and a melanoma antigen, where the melanoma antigen
CC   is selected from tyrosinase, tyrosinase related protein 1, tyrosinase
CC   related protein 2, gp 100, MAGE antigens, BAGE antigens, NYEs01, MART
CC   antigens, GM2, antigenic portions and combinations of these. The melanoma
CC   antigen is covalently bound to a javelin molecule, where the melanoma
CC   antigen bound to the javelin molecule is non-covalently bound to the heat
CC   shock protein. The composition is useful for inducing an immune response
CC   for the treatment of melanoma. AAU71980-AAU72481 represent melanoma
CC   antigen peptides of the invention
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 45;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              |||||||||
Db          1 IMDQVPFSV 9


RESULT 4
AAU10221
ID   AAU10221 standard; peptide; 9 AA.
XX
AC   AAU10221;
XX
DT   27-FEB-2002  (first entry)
XX
DE   Cancer cell associated peptide G9-209-2M.
XX
KW   Human; cancer cell associated peptide; G9-209-2M; CTL;
KW   single chain major histocompatibility complex class I; MHC;
KW   human beta-2 microglobulin; cytotoxic T lymphocyte; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200172768-A2.
XX
```

PD 04-OCT-2001.
XX
PF 19-MAR-2001; 2001WO-IL000260.
XX
PR 27-MAR-2000; 2000US-00534966.
XX
PA (TECR) TECHNION RES & DEV FOUND LTD.
XX
PI Reiter Y;
XX
DR WPI; 2001-656911/75.
XX
PT New major histocompatibility complex (MHC) class I polypeptide, useful
PT for presenting antigenic peptides to cytotoxic T lymphocyte clones,
PT comprises beta-2 microglobulin covalently linked to MHC class I heavy
PT chain.
XX
PS Example 1; Page 27; 81pp; English.
XX
CC The invention relates to a recombinant single chain major
CC histocompatibility complex (MHC) class I polypeptide comprising an amino
CC acid sequence including a functional human beta-2 microglobulin
CC (in)directly covalently linked to a functional human MHC class I heavy
CC chain and the nucleic acids encoding it. The recombinant MHC polypeptide
CC is used in a method is for generating large quantities of pure single
CC chain MHC class I polypeptides, which is useful in monomeric or
CC multimeric forms to present antigenic peptides to cytotoxic T lymphocyte
CC (CTL) clones. The present sequence is a cancer cell associated peptide
CC which can bind MHC in an experiment demonstrating the production of the
CC recombinant MHC complexes
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 45; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IMDQVPFSV 9
|||||||||
Db 1 IMDQVPFSV 9

RESULT 5
AAE05120
ID AAE05120 standard; peptide; 9 AA.
XX
AC AAE05120;
XX
DT 18-SEP-2001 (first entry)
XX
DE Modified tumour-associated antigen, GP100 peptide, CLP 572.
XX
KW Tumour-associated antigen; TAA; GP100 antigen; cytostatic; gene therapy;
KW immune response; tetanus toxoid; TT; diphtheria toxoid; DT; prophylactic;
KW vaccine; cancer; therapeutic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 2
FT /note= "Wild type Thr substituted with Met"
XX
PN WO200149317-A2.
XX
PD 12-JUL-2001.
XX
PF 05-JAN-2001; 2001WO-CA000005.
XX
PR 05-JAN-2000; 2000US-0174587P.
XX
PA (AVET) AVENTIS PASTEUR LTD.
XX
PI Emtage P, Barber BH, Sambhara S, Sia CDY;
XX
DR WPI; 2001-441790/47.
XX
PT Enhancing immune response to antigen such as tumor antigen for treating
PT cancer in an animal involves administering an inducing agent to the
PT animal followed by administering inducing agent-antigen mixture.
XX
PS Example 1; Page 28; 62pp; English.
XX
CC The invention relates to a method of enhancing an immune response against
CC tumour-associated antigens (TAAs), such as GP100 and carcinoembryonic
CC antigen (CEA) in an animal. The method involves priming of the animal
CC with an inducing agent such as tetanus toxoid (TT) or diphtheria toxoid
CC (DT), subsequently followed by administration of an inducing agent-
CC antigen mixture. The method provides the enhancement or augmentation of
CC the immune response to the antigen and/or improves a vaccination protocol
CC by allowing use of less antigen. The immunisation of the animal with
CC tumour-associated antigen is useful for the prophylactic or therapeutic
CC treatment of cancer. The present sequence is modified tumour-associated
CC antigen, GP100 peptide fragment related to the invention
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 45; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IMDQVPFSV 9
|||||||||
Db 1 IMDQVPFSV 9

RESULT 6
AAO17085
ID AAO17085 standard; peptide; 9 AA.
XX
AC AAO17085;
XX
DT 06-JUN-2002 (first entry)
XX
DE Gp 100 analogue antigen SEQ ID NO: 5.
XX
KW Cryopreserved mature dendritic cell; antigen; vaccine; cytostatic;
KW virucide; cancer; hepatitis B virus.
XX
OS Unidentified.
XX
PN WO200216560-A1.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-EP009790.
XX
PR 24-AUG-2000; 2000DE-01041515.
XX
PA (SCHU/) SCHULER G.
XX
PI Schuler G, Schuler-Thurner B;
XX
DR WPI; 2002-292062/33.
XX
PT Preparation of cryopreserved, mature dendritic cells, useful in vaccines,
PT comprises culturing immature cells on medium containing cocktail of
PT maturation factors, then freezing.
XX
PS Disclosure; Fig 28; 87pp; German.
XX
CC The present invention relates to a method for the preparation of ready-
CC for-use, cryopreserved, mature dendritic cells comprising growing
CC immature dendritic cells in a culture medium that includes a 'maturation
CC cocktail' of one or more maturation stimuli and freezing the resulting
CC matured cells in a freezing medium that does not contain heterologous
CC serum. When loaded with antigens, the dendritic cells can be used as

```
CC   vaccines, e.g. against tumours and hepatitis B virus. The present
CC   sequence is an antigen described in the invention
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 45;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              |||||||||
Db          1 IMDQVPFSV 9


RESULT 7
AAE17297
ID   AAE17297 standard; peptide; 9 AA.
XX
AC   AAE17297;
XX
DT   18-APR-2002  (first entry)
XX
DE   gp100-modified peptide.
XX
KW   Artificial antigen presenting cell; AAPC; beta2-microglobulin;
KW   human leukocyte antigen; HLA; major histocompatibility complex; MHC;
KW   cytotoxic T lymphocyte; CTL; T cell-specific antigen; TCA; antitumour;
KW   immune response; cancer.
XX
OS   Synthetic.
XX
PN   WO200194944-A2.
XX
PD   13-DEC-2001.
XX
PF   01-JUN-2001; 2001WO-US017981.
XX
PR   02-JUN-2000; 2000US-0209157P.
XX
PA   (SLOK ) SLOAN KETTERING INST CANCER RES.
XX
PI   Sadelain M,  Latouche J;
XX
DR   WPI; 2002-139667/18.
XX
PT   Artificial antigen presenting cells for activating T lymphocytes,
PT   comprises eukaryotic cell expressing antigen presenting complex having
PT   beta2-microglobulin, exogenous accessory molecule, human leukocyte
PT   antigen molecule and protein.
XX
PS   Example 7; Page 29; 75pp; English.
XX
CC   The present invention relates to an artificial antigen presenting cell
CC   (AAPC) comprising a eukaryotic cell expressing an antigen presenting
CC   complex comprising beta2-microglobulin, an exogenous accessory molecule,
CC   a human leukocyte antigen, HLA (major histocompatibility complex, MHC)
CC   molecule of a single type and a protein that is processed intracellularly
CC   to produce an exogenous T cell-specific epitope. The invention also
CC   relates to methods for activation of T lymphocytes. The method is also
CC   useful for identifying within a test population of cytotoxic T
CC   lymphocytes (CTLs), CTLs specifically activated against a known T-cell
CC   specific antigen (TCA), which is useful for diagnostic purposes. AAPC is
CC   also useful for activating CTLs, by contacting AAPC with a suitable
CC   population of T lymphocytes under conditions suitable for the activation
CC   and isolating the activated CTLs. AAPC is further useful for the
CC   investigation of primary T cell activation and diagnostic applications
CC   here primary T cell activation allow discovery of antigens and accessory
CC   molecules, and diagnostic applications include cell-based assays for
CC   quantifying immune response in normal, infected or treated (vaccinated)
CC   patients. Composition comprising AAPC or activated T cells produced by
CC   utilising AAPC is useful for eliciting an antitumour response. The
CC   invention is used for the treatment of cancer. The present sequence is gp
CC   -100 modified peptide used in the exemplification of the invention
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 45;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              |||||||||
Db          1 IMDQVPFSV 9


RESULT 8
AAE36055
ID   AAE36055 standard; peptide; 9 AA.
XX
AC   AAE36055;
XX
DT   26-JUN-2003  (first entry)
XX
DE   Tumour associated HLA-A2-restricted peptide, G9-209-2M.
XX
KW   Major histocompatibility complex; MHC; immune deception; cytostatic;
KW   therapy; cytotoxic T cell; immunostimulant; cancer.
XX
OS   Unidentified.
XX
PN   WO2002102299-A2.
XX
PD   27-DEC-2002.
XX
PF   18-JUN-2002; 2002WO-IL000478.
XX
PR   19-JUN-2001; 2001US-0298915P.
PR   29-MAR-2002; 2002US-00108511.
XX
PA   (TECR ) TECHNION RES & DEV FOUND LTD.
XX
PI   Reiter Y,  Lev A;
XX
DR   WPI; 2003-210086/20.
XX
PT   New immuno-molecules comprising a soluble human major histocompatibility
PT   complex class I effector domain and a targeting domain linked to the
PT   effector domain, useful for immune deception, particularly in treating
PT   cancer.
XX
PS   Example; Page 50; 51pp; English.
XX
CC   The invention relates to immuno-molecules comprising a soluble human
CC   major histocompatibility complex (MHC) class I effector domain and a
CC   targetting domain linked to the effector domain, useful for immune
CC   deception. The invention also relates to methods for immune deception.
CC   The method is useful for producing an immunomolecule, and selectively
CC   killing a cell in a patient. The immuno-molecules are useful for immune
CC   deception, particularly treating cancer. The immuno-molecules and methods
CC   are useful for recruiting active cytotoxic T cells for tumour killing via
CC   cancer-specific antibody or ligand guided targetting of single-chain MHC-
CC   peptide complexes. The present sequence is a tumour associated HLA-A2-
CC   restricted peptide used to illustrate the method of the invention
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 45;  DB 6;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              |||||||||
Db          1 IMDQVPFSV 9


RESULT 9
```

AAG79857
ID AAG79857 standard; peptide; 9 AA.
XX
AC AAG79857;
XX
DT 28-APR-2003 (first entry)
XX
DE HLA-A2-restricted epitope of the gp100 melanoma antigen.
XX
KW Epitope; dendritic cell; tumour antigen-specific; CD8+ human cell clone;
KW M15; HLA-A2-restricted epitope; gp100 melanoma antigen; M99;
KW HLA-B7-restricted CD8+ CTL clone; carcinoembryonic antigen; cancer;
KW haematological cancer; neurological cancer; melanoma; breast cancer;
KW lung cancer; head; neck; gastrointestinal cancer; genitourinary cancer;
KW bone cancer; vascular cancer; infection; influenza virus; HIV;
KW M. tuberculosis; P. falciparum.
XX
OS Homo sapiens.
XX
PN WO200286083-A2.
XX
PD 31-OCT-2002.
XX
PF 22-APR-2002; 2002WO-US012733.
XX
PR 20-APR-2001; 2001US-0285137P.
XX
PA (MAYO-) MAYO FOUND MEDICAL EDUCATION RES.
XX
PI Chen L, Strome SE;
XX
DR WPI; 2003-067637/06.
XX
PT Enhancing T cell responsiveness in a mammal, useful for treating cancer
PT or other infections, comprises administering to the subject a compound
PT comprising an agent that interferes with an interaction between B7-H1 and
PT a T cell.
XX
PS Example 1; Page 31; 55pp; English.
XX
CC The sequences given in AAG79857-58 are epitopes which were used to
CC stimulate dendritic cells to generate tumour antigen-specific CD8+ human
CC cell clones. M15 is a human CD8+ CTL clone that specifically recognises
CC an HLA-A2-restricted epitope of the gp100 melanoma antigen. M99 is an HLA
CC -B7-restricted CD8+ CTL clone which recognises an epitope pf
CC carcinoembryonic antigen. The human T cell clones which are generated
CC were used in the method of the invention for treating cancer, such as
CC haematological cancer, neurological cancer, melanoma, breast cancer, lung
CC cancer, head and neck cancer, gastrointestinal cancer, genitourinary
CC cancer, bone cancer or vascular cancer, or infections by influenza virus,
CC HIV, M. tuberculosis, or P. falciparum
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 45; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IMDQVPFSV 9
|||||||||
Db 1 IMDQVPFSV 9

RESULT 10
AAE35575
ID AAE35575 standard; peptide; 9 AA.
XX
AC AAE35575;
XX
DT 17-JUN-2003 (first entry)
XX
DE Melanoma gp100 eptiope.
XX
KW Fusion agent; immunogenic; proliferative disease; infectious disease;
KW cancer; therapy; vaccine; melanoma; Trojan antigen; TA; epitope.
XX
OS Unidentified.
XX
PN WO200294994-A2.
XX
PD 28-NOV-2002.
XX
PF 20-MAY-2002; 2002WO-US015992.
XX
PR 18-MAY-2001; 2001US-0291874P.
XX
PA (MAYO-) MAYO FOUND MEDICAL EDUCATION RES.
XX
PI Celis E;
XX
DR WPI; 2003-140367/13.
XX
PT Fusion agent useful for preventing and treating an infectious disease, or
PT a proliferative disease, such as cancer, comprises a transport domain,
PT two cleavage sites, a peptide epitope and a biologically active agent.
XX
PS Example 1; Page 37; 72pp; English.
XX
CC The invention relates to a fusion agent (Trojan antigen; TA) comprising a
CC transport domain, two cleavage sites, a peptide epitope recognised by an
CC antigen-specific receptor on an effector T-lymphocyte precursor cell and
CC a biologically active agent, where there is a cleavage site between the
CC peptide epitope and the biologically active agent and between each
CC biologically active agent. The fusion agent is used to make a cell
CC immunogenic or antigenic. It is also useful for preventing and treating
CC an infectious disease such as viral, bacterial, protozoal, fungal or
CC yeast disease, or proliferative disease such as cancer (e.g. melanoma,
CC neural tissue, gastrointestinal, breast, lung, ovarian, testicular,
CC prostate, cervical, bladder, vaginal, liver, renal, bone, haematological
CC or vascular tissue cancer). The invention is used as vaccines. The
CC present sequence is melanoma gp100 eptiope. This peptide is used in the
CC exemplification of the invention
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 45; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IMDQVPFSV 9
|||||||||
Db 1 IMDQVPFSV 9

RESULT 11
ABJ19877
ID ABJ19877 standard; peptide; 9 AA.
XX
AC ABJ19877;
XX
DT 10-APR-2003 (first entry)
XX
DE MHC binding peptide SEQ ID No 42.
XX
KW Antirheumatic; antiallergic; antiarthritic; nootropic; neuroprotective;
KW antiinflammatory; major histocompatibility complex; MHC;
KW autoimmune disease; T cell; B cell; allergic disease; multiple sclerosis;
KW rheumatoid arthritis; neurodegenerative disorder; Alzheimer's disease;
KW inflammation; gene therapy; MHC binding peptide.
XX
OS Synthetic.
XX
PN WO200294981-A2.
XX
PD 28-NOV-2002.
XX

PF 16-MAY-2002; 2002WO-IL000383.
XX
PR 16-MAY-2001; 2001US-0290958P.
PR 29-MAY-2001; 2001US-00865548.
XX
PA (TECR) TECHNION RES & DEV FOUND LTD.
XX
PI Barnea E, Beer I, Ziv T, Admon A, Dassau L, Buchsbaum S;
XX
DR WPI; 2003-210043/20.
XX
PT Identifying peptides that are capable of binding to major
PT histocompatibility complex (MHC) molecules of a particular haplotype by
PT analyzing peptides bound to the soluble and secreted form of the MHC
PT molecules of the particular haplotype.
XX
PS Example; Page 156; 238pp; English.
XX
CC The invention relates to a novel method for identifying peptides
CC originating from a particular cell type, which are capable of binding to
CC major histocompatibility complex (MHC) molecules of a particular
CC haplotype. The method comprises analysing peptides bound to the soluble
CC and secreted form of the MHC molecules of the particular haplotype. The
CC method is useful for identifying peptides for treating an autoimmune
CC disease, such as T or B cell and/or allergic disease or condition,
CC rheumatoid arthritis, or multiple sclerosis, neurodegenerative disorders,
CC e.g. Alzheimer's disease, or diseases associated with inflammation. The
CC sequences of the invention may be used in a gene therapy application.
CC This sequence represents a peptide relating to the method for identifying
CC MHC binding peptides of the invention
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 45; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IMDQVPFSV 9
|||||||||
Db 1 IMDQVPFSV 9

RESULT 12
ABU08668
ID ABU08668 standard; peptide; 9 AA.
XX
AC ABU08668;
XX
DT 10-JUN-2003 (first entry)
XX
DE Cancer cell associated peptide G9-209-2M.
XX
KW Chimeric polypeptide; antigenic peptide; MHC class I;
KW major histocompatibility complex class I; beta-2 microglobulin;
KW MHC class I heavy chain; single chain MHC class I;
KW CTL: cytotoxic T-lymphocyte; MHC binding T cell epitope;
KW primary CTL induction; melanoma antigen gp100; cancer; G9-209-2M.
XX
OS Synthetic.
XX
PN US2003003535-A1.
XX
PD 02-JAN-2003.
XX
PF 13-FEB-2002; 2002US-00073300.
XX
PR 27-MAR-2000; 2000US-00534966.
XX
PA (TECR) TECHNION RES & DEV FOUND LTD.
XX
PI Reiter Y;
XX
DR WPI; 2003-352830/33.
XX
PT New chimeric polypeptide useful for generating antibodies, comprises an
PT antigenic peptide that binds a human major histocompatibility complex
PT (MHC) class I, a human beta-2 microglobulin and a human MHC class I heavy
PT chain.
XX
PS Example 1; Page 9; 37pp; English.
XX
CC The invention describes a chimeric polypeptide (I) comprising an
CC antigenic peptide capable of binding a human major histocompatibility
CC complex (MHC) class I, a functional human beta-2 microglobulin and a
CC functional human MHC class I heavy chain. The chimeric polypeptide is
CC useful in generating large quantities of pure single chain MHC class I
CC polypeptides which can be used in monomeric or multimeric form to present
CC antigenic peptides to cytotoxic T-lymphocyte (CTL) clones. They may also
CC be used in rapid, sensitive and reliable MHC peptide binding assay to
CC identify high affinity MHC binding T cell epitopes, in in vitro primary
CC CTL induction studies to define those peptides that are immunogenic, and
CC to generate antibodies. This is the amino acid sequence of a cancer cell
CC associated peptide derived from melanoma antigen gp100 and used for major
CC histocompatibility (MHC) binding
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 45; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IMDQVPFSV 9
|||||||||
Db 1 IMDQVPFSV 9

RESULT 13
ADA89157
ID ADA89157 standard; peptide; 9 AA.
XX
AC ADA89157;
XX
DT 20-NOV-2003 (first entry)
XX
DE Gp100-derived peptide G9-209 SEQ ID NO:1.
XX
KW immunoglobulin; Ig; heavy chain variable domain;
KW light chain variable domain; major histocompatibility complex; MHC;
KW gp100; MUC1; TAX; hTERT; cytostatic; gene therapy; cancerous disorder;
KW cancer; gene; ds.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO2003070752-A2.
XX
PD 28-AUG-2003.
XX
PF 20-FEB-2003; 2003WO-US005128.
XX
PR 20-FEB-2002; 2002US-0358994P.
XX
PA (DYAX-) DYAX CORP.
PA (TECR) TECHNION RES & DEV FOUND LTD.
XX
PI Hoogenboom HRJM, Reiter Y;
XX
DR WPI; 2003-663847/62.
XX
PT New protein comprising an immunoglobulin heavy chain variable (VH) domain
PT and an immunoglobulin light chain variable (VL) domain, useful for
PT preparing a composition for treating or preventing a cancerous disorder.
XX
PS Claim 4; Page 137; 224pp; English.
XX
CC The present invention describes a protein comprising an immunoglobulin

CC (Ig) heavy chain variable (VH) domain and an Ig light chain variable (VL)
CC domain. The protein binds a complex comprising a major histocompatibility
CC complex (MHC) and a peptide, does not substantially bind the MHC in the
CC absence of the bound peptide, and does not substantially bind the peptide
CC in the absence of the MHC. The peptide is a peptide fragment of gp100,
CC MUC1, TAX or hTERT. ALso described: (1) a pharmaceutical composition
CC comprising the novel protein and a carrier; (2) a cytotoxic T cell
CC comprising one or more nucleic acids for expressing the Ig that binds a
CC complex having an MHC and a peptide, does not substantially bind the MHC
CC in the absence of the bound peptide, and does not substantially bind the
CC peptide in the absence of the MHC; (3) an isolated nucleic acid
CC comprising a first segment that encodes the Ig variable domain; (4) a
CC host cell comprising heterologous nucleic acid sequences that encodes the
CC novel protein; (5) a transgenic animal whose genome includes heterologous
CC nucleic acid sequences that encode the protein; (6) identifying the
CC protein that specifically binds the MHC-peptide complex; (7) expressing
CC an antigen-binding protein; (8) ablating or killing a target cell that
CC displays a peptide on a surface MHC molecule; (9) treating or preventing
CC a cancerous disorder in a subject; and (10) detecting an MHC-peptide
CC complex in a sample. A protein of the invention has cytostatic activity,
CC and can be used in gene therapy. The protein is useful for preparing a
CC composition for treating or preventing a cancerous disorder. The present
CC sequence represents a gp100-derived peptide, which is used in the
CC exemplification of the present invention.
XX
SQ Sequence 9 AA;

```
  Query Match             100.0%;  Score 45;  DB 6;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              |||||||||
Db          1 IMDQVPFSV 9
```

RESULT 14
AAE38651
ID AAE38651 standard; peptide; 9 AA.
XX
AC AAE38651;
XX
DT 04-DEC-2003 (first entry)
XX
DE HLA-A2 restricted peptide, gp100 (209).
XX
KW Major histocompatibility complex; MHC; HLA-restricted antigen; cancer;
KW viral infection; autoimmune disease; gene therapy; cytostatic; virucide;
KW immunomodulator.
XX
OS Unidentified.
XX
PN WO2003068201-A2.
XX
PD 21-AUG-2003.
XX
PF 11-FEB-2003; 2003WO-IL000105.
XX
PR 13-FEB-2002; 2002US-00073301.
XX
PA (TECR) TECHNION RES & DEV FOUND LTD.
XX
PI Reiter Y, Denkberg G;
XX
DR WPI; 2003-689603/65.
XX
PT New isolated molecule comprising an antibody that binds with a human
PT major histocompatibility complex (MHC) class I being complexed with a HLA
PT -restricted antigen, useful for treating cancer, viral infection or
PT autoimmune disease.
XX
PS Example; Page 10; 81pp; English.
XX
CC The invention relates to an isolated molecule comprising an antibody
CC specifically bindable with a binding affinity below 20 nanomolar to a
CC human major histocompatibility complex (MHC) class I being complexed with
CC a HLA-restricted antigen. The molecules, antibodies, and methods are
CC useful for treating cancer, viral infection and an autoimmune disease.
CC The invention is useful in gene therapy. The present sequence is HLA-A2
CC restricted peptide
XX
SQ Sequence 9 AA;

```
  Query Match             100.0%;  Score 45;  DB 7;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              |||||||||
Db          1 IMDQVPFSV 9
```

RESULT 15
ADB97733
ID ADB97733 standard; peptide; 9 AA.
XX
AC ADB97733;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human gp100 antigenic peptide #12 analogue #2.
XX
KW Human; melanoma antigen recognised by T-lymphocytes; MART-1; melanoma;
KW skin cancer; T lymphocyte; cytostatic; gene therapy; vaccine; antigen;
KW major histocompatibility complex; MHC; human leukocyte antigen; HLA-A2;
KW tumour infiltrating lymphocyte; mutant; mutein.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN US2003144482-A1.
XX
PD 31-JUL-2003.
XX
PF 03-JUL-2001; 2001US-00898860.
XX
PR 22-APR-1994; 94US-00231565.
PR 12-MAR-1999; 99US-00267439.
XX
PA (KAWA/) KAWAKAMI Y.
PA (ROSE/) ROSENBERG S A.
XX
PI Kawakami Y, Rosenberg SA;
XX
DR WPI; 2003-755536/71.
XX
PT New immunogenic peptides derived from melanoma antigens recognized by T-
PT lymphocytes or from gp100, useful for preventing or treating melanoma.
XX
PS Example 5; Page 31; 77pp; English.
XX
CC The invention relates to an immunogenic peptide having contiguous amino
CC acids derived from the sequence of melanoma antigens recognised by T-
CC lymphocytes (MART-1) or gp100. The MART-1 sequence appears as ADB97651,
CC and the gp100 (differing by 1 amino acid from the previously published
CC gp100 (ADB97770) appears as ADB97676. Also included are a pharmaceutical
CC composition (comprising the above peptide and an excipient, diluent or
CC carrier), a vaccine for immunising a mammal (comprising the above peptide
CC in a carrier), preventing or treating melanoma (comprising administering
CC the above composition to a mammal in an amount to stimulate the
CC production of protective antibodies or immune cells), a purified and
CC isolated nucleic acid sequence encoding the above peptide, a recombinant
CC expression vector comprising at least one nucleic acid sequence cited
CC above, a host organism transformed or transfected with the vector
CC (expressing the peptide) and antibodies reactive with the above
CC immunogenic peptide. The peptide sequence contains at least one amino

CC acid modification (amino acid substitution)of the MART-1 or gp100
CC sequence to enhance binding of the peptide to a Major Histocompatibility
CC Complex (MHC) molecule. The peptide is recognised by Human Leukocyte
CC Antigen (HLA)-A2 restricted tumour infiltrating lymphocytes. The
CC composition and methods are useful in preventing or treating melanoma and
CC skin cancer. The present sequence represents a modified melanoma
CC antigenic peptide of the invention.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 45; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 IMDQVPFSV 9
              |||||||||
Db          1 IMDQVPFSV 9
```

Search completed: December 8, 2004, 17:16:41
Job time : 11.2927 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2004, 17:06:18 ; Search time 1.7561 Seconds
(without alignments)
493.111 Million cell updates/sec

Title: US-10-073-301A-2
Perfect score: 45
Sequence: 1 IMDQVPFSV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	88.9	274	2	T16581	hypothetical prote
2	39	86.7	491	2	A49179	melanoma antigen h
3	39	86.7	626	2	S53871	Pmel 17 protein -
4	39	86.7	662	2	I38400	melanoma-associate
5	39	86.7	668	2	A41234	melanocyte-specifi
6	33	73.3	350	2	I78848	LIMK-2c - rat
7	33	73.3	439	2	E81291	hypothetical prote
8	33	73.3	617	1	I78847	LIM motif-containi
9	33	73.3	617	2	JC5814	LIM motif-containi
10	33	73.3	638	1	I78846	LIM motif-containi
11	33	73.3	638	2	JC5813	LIM motif-containi
12	32	71.1	128	2	E46178	probable transcrip
13	32	71.1	197	2	S65051	low molecular weig
14	32	71.1	197	2	S72398	low molecular weig
15	32	71.1	249	2	C90433	ABC transporter, A
16	32	71.1	258	2	T51689	probable transcrip
17	32	71.1	288	2	JC5262	leucine zipper pro
18	32	71.1	365	2	D86470	F21H2.9 protein -
19	32	71.1	633	2	E97999	hypothetical prote
20	32	71.1	633	2	H95128	ABC transporter, A
21	32	71.1	791	2	A53691	diacylglycerol kin
22	31	68.9	126	2	T22189	hypothetical prote
23	31	68.9	199	2	F71462	hypothetical prote
24	31	68.9	258	2	B97187	polysaccharide ABC
25	31	68.9	363	2	T19032	hypothetical prote
26	31	68.9	381	2	T13666	NADH2 dehydrogenas
27	31	68.9	609	2	T48762	cytochrome p450 re
28	31	68.9	664	2	T35122	probable ATP-depen
29	31	68.9	770	2	S00643	anthranilate synth
30	31	68.9	770	2	S11161	anthranilate synth
31	31	68.9	837	2	I57557	DNA-Binding Protei
32	31	68.9	848	2	A54740	interleukin-4-indu
33	30	66.7	126	2	S77183	hypothetical prote
34	30	66.7	156	2	I49446	8-oxo-dGTPase - mo
35	30	66.7	200	2	E81314	probable periplasm
36	30	66.7	248	2	B83134	probable pili asem
37	30	66.7	263	2	E75046	hypothetical prote
38	30	66.7	319	2	D97205	hypothetical prote
39	30	66.7	388	2	D75353	yellow-related pro
40	30	66.7	389	2	H90413	conserved hypothet
41	30	66.7	390	2	AI0396	multidrug resistan
42	30	66.7	410	2	T38815	septin homolog spn
43	30	66.7	455	2	T32189	zinc finger protei
44	30	66.7	474	2	T29336	hypothetical prote
45	30	66.7	482	2	T43996	virion protein [im

ALIGNMENTS

RESULT 1
T16581
hypothetical protein K08A8.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T16581
R;Pauley, A.
submitted to the EMBL Data Library, October 1995
A;Description: The sequence of C. elegans cosmid K08A8.
A;Reference number: Z18541
A;Accession: T16581
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-274 <PAU>
A;Cross-references: UNIPROT:Q21305; EMBL:U38377; NID:g1022968; PID:g1022970; PIDN:AAA79
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:K08A8.2
A;Introns: 50/2; 71/3; 113/3; 148/3

Query Match 88.9%; Score 40; DB 2; Length 274;
Best Local Similarity 66.7%; Pred. No. 1.1;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 IMDQVPFSV 9
              :|||:|||:
Db        176 VMDQIPFSL 184
```

RESULT 2
A49179
melanoma antigen homolog rpe1 - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A49179; I45861
R;Kim, R.Y.; Wistow, G.J.
Exp. Eye Res. 55, 657-662, 1992
A;Title: The cDNA RPE1 and monoclonal antibody HMB-50 define gene products preferential
A;Reference number: A49179; MUID:93122163; PMID:1478275
A;Accession: A49179
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-491 <KIM>
A;Cross-references: UNIPROT:Q06154
A;Experimental source: retinal pigment epithelium
A;Note: sequence extracted from NCBI backbone (NCBIN:122438, NCBIP:122439)
C;Genetics:
A;Gene: RPE1

Query Match 86.7%; Score 39; DB 2; Length 491;
Best Local Similarity 88.9%; Pred. No. 3.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 IMDQVPFSV 9
              | |||||||
Db         52 ITDQVPFSV 60
```

RESULT 3
S53871
Pmel 17 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S53871
R;Kwon, B.S.; Halaban, R.; Ponnazhagan, S.; Kim, K.; Chintamaneni, C.; Bennett, D.; Pick
Nucleic Acids Res. 23, 154-158, 1995
A;Title: Mouse silver mutation is caused by a single base insertion in the putative cyto
A;Reference number: S53871; MUID:95175358; PMID:7870580
A;Accession: S53871
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-626 <KWO>
A;Cross-references: UNIPROT:Q60696; GB:U14133; NID:g887940; PIDN:AAA69538.1; PID:g887941

```
  Query Match             86.7%;  Score 39;  DB 2;  Length 626;
  Best Local Similarity   88.9%;  Pred. No. 4.6;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              | |||||||
Db        209 ITDQVPFSV 217
```

RESULT 4
I38400
melanoma-associated ME20 antigen (me20m) - human
N;Alternate names: melanoma antigen 25
C;Species: Homo sapiens (man)
C;Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C;Accession: I38400; A53668; A55753
R;Maresh, G.A.; Marken, J.S.; Neubauer, M.; Aruffo, A.; Hellstrom, I.; Hellstrom, K.; Ma
DNA Cell Biol. 13, 87-95, 1994
A;Title: Cloning and expression of the gene for the Melanoma-Associated ME20 Antigen.
A;Reference number: I38400; MUID:94235165; PMID:8179825
A;Accession: I38400
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-662 <RES>
A;Cross-references: UNIPROT:P40967; EMBL:U01874; NID:g494939; PIDN:AAA18479.1; PID:g4949
R;Adema, G.J.; de Boer, A.J.; Vogel, A.M.; Loenen, W.A.M.; Figdor, C.G.
J. Biol. Chem. 269, 20126-20133, 1994
A;Title: Molecular characterization of the melanocyte lineage-specific antigen gp100.
A;Reference number: A53668; MUID:94327568; PMID:7519602
A;Accession: A53668
A;Molecule type: mRNA
A;Residues: 1-592,594-662 <ADE>
A;Cross-references: GB:S73003; NID:g639589; PIDN:AAC60634.1; PID:g639590
R;Kawakami, Y.; Eliyahu, S.; Delgado, C.H.; Robbins, P.F.; Sakaguchi, K.; Appella, E.; Y
Proc. Natl. Acad. Sci. U.S.A. 91, 6458-6462, 1994
A;Title: Identification of a human melanoma antigen recognized by tumor-infiltrating lym
A;Reference number: A55753; MUID:94294401; PMID:8022805
A;Accession: A55753
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-161,'F',163-592,594-662 <KAW>
C;Keywords: glycoprotein

```
  Query Match             86.7%;  Score 39;  DB 2;  Length 662;
  Best Local Similarity   88.9%;  Pred. No. 4.9;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              | |||||||
Db        209 ITDQVPFSV 217
```

RESULT 5
A41234
melanocyte-specific protein Pmel-17 precursor - human
C;Species: Homo sapiens (man)
C;Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 30-Sep-1993
C;Accession: A41234
R;Kwon, B.S.; Chintamaneni, C.; Kozak, C.A.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.,
Proc. Natl. Acad. Sci. U.S.A. 88, 9228-9232, 1991
A;Title: A melanocyte-specific gene, Pmel 17, maps near the silver coat color locus on
A;Reference number: A41234; MUID:92021023; PMID:1924386
A;Accession: A41234
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-668 <KWO>
A;Cross-references: GB:M77348

```
  Query Match             86.7%;  Score 39;  DB 2;  Length 668;
  Best Local Similarity   88.9%;  Pred. No. 4.9;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              | |||||||
Db        209 ITDQVPFSV 217
```

RESULT 6
I78848
LIMK-2c - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004
C;Accession: I78848
R;Nunoue, K.; Ohashi, K.; Okano, I.; Mizuno, K.
Oncogene 11, 701-710, 1995
A;Title: LIMK-1 and LIMK-2, two members of a LIM motif-containing protein kinase family
A;Reference number: I58353; MUID:95380177; PMID:7651734
A;Accession: I78848
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-350 <RES>
A;Cross-references: UNIPROT:P53670; GB:D31876; NID:g1684614; PIDN:BAA06675.1; PID:g1000
F;51-103/Domain: LIM metal-binding repeat homology <LIM2>

```
  Query Match             73.3%;  Score 33;  DB 2;  Length 350;
  Best Local Similarity   55.6%;  Pred. No. 44;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              : ||:|:||
Db        121 VQDQLPYSV 129
```

RESULT 7
E81291
hypothetical protein Cj1454c [imported] - Campylobacter jejuni (strain NCTC 11168)
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: E81291
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilli
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barre
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy
A;Reference number: A81250; MUID:20150912; PMID:10688204
A;Accession: E81291
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-439 <PAR>
A;Cross-references: UNIPROT:Q9PMK6; GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB738
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: Cj1454c
C;Superfamily: conserved hypothetical protein b0835

Query Match 73.3%; Score 33; DB 2; Length 439;
Best Local Similarity 75.0%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MDQVPFSV 9
|:|||| |
Db 335 MEQVPFKV 342

RESULT 8
I78847
LIM motif-containing protein kinase (EC 2.7.1.-) 2, splice form b [similarity] - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 16-Jun-2000 #sequence_revision 16-Jun-2000 #text_change 09-Jul-2004
C;Accession: I78847; I78849
R;Nunoue, K.; Ohashi, K.; Okano, I.; Mizuno, K.
Oncogene 11, 701-710, 1995
A;Title: LIMK-1 and LIMK-2, two members of a LIM motif-containing protein kinase family.
A;Reference number: I58353; MUID:95380177; PMID:7651734
A;Accession: I78847
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-617 <RES>
A;Cross-references: UNIPROT:P53670; GB:D31875; NID:g1684613; PIDN:BAA06674.1; PID:g10006
A;Accession: I78849
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-163 <RE2>
A;Cross-references: GB:D31877; NID:g1684615; PIDN:BAA06676.1; PID:g1000690
C;Genetics: <RES1>
A;Gene: limk-2b
C;Genetics: <RES2>
A;Gene: limk-2d
C;Superfamily: LIM protein kinase; LIM metal-binding repeat homology; protein kinase hom
C;Keywords: alternative splicing; ATP; duplication; phosphotransferase; serine/threonine
F;51-103/Domain: LIM metal-binding repeat homology <LIM2>
F;308-587/Domain: protein kinase homology <KIN>
F;316-324/Region: protein kinase ATP-binding motif

Query Match 73.3%; Score 33; DB 1; Length 617;
Best Local Similarity 55.6%; Pred. No. 82;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IMDQVPFSV 9
: ||:|:||
Db 121 VQDQLPYSV 129

RESULT 9
JC5814
LIM motif-containing protein kinase (EC 2.7.1.-) 2, splice form b - mouse
C;Species: Mus musculus (house mouse)
C;Date: 04-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C;Accession: JC5814
R;Koshimizu, U.; Takahashi, H.; Yoshida, M.C.; Nakamura, T.
Biochem. Biophys. Res. Commun. 241, 243-250, 1997
A;Title: cDNA cloning, genomic organization, and chromosomal localization of the mouse l
A;Reference number: JC5813; MUID:98086337; PMID:9425257
A;Accession: JC5814
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-617 <KOS>
A;Cross-references: UNIPROT:O54785; DDBJ:AB008119
A;Experimental source: embryo
C;Comment: This enzyme is involved in a signal transduction pathway to regulate various
C;Genetics:
A;Gene: Limk2b
A;Map position: 1d
C;Superfamily: LIM protein kinase; LIM metal-binding repeat homology; protein kinase hom
C;Keywords: alternative splicing; phosphotransferase
F;51-103/Domain: LIM metal-binding repeat homology <LIM2>
F;308-587/Domain: protein kinase homology <KIN>
F;317-580/Domain: protein kinase #status predicted <PKD>

Query Match 73.3%; Score 33; DB 2; Length 617;
Best Local Similarity 55.6%; Pred. No. 82;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IMDQVPFSV 9
: ||:|:||
Db 121 VQDQLPYSV 129

RESULT 10
I78846
LIM motif-containing protein kinase (EC 2.7.1.-) 2, splice form a [similarity] - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 16-Jun-2000 #sequence_revision 16-Jun-2000 #text_change 09-Jul-2004
C;Accession: I78846
R;Nunoue, K.; Ohashi, K.; Okano, I.; Mizuno, K.
Oncogene 11, 701-710, 1995
A;Title: LIMK-1 and LIMK-2, two members of a LIM motif-containing protein kinase family
A;Reference number: I58353; MUID:95380177; PMID:7651734
A;Accession: I78846
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-638 <RES>
A;Cross-references: UNIPROT:P53670; GB:D31874; NID:g1684612; PIDN:BAA06673.1; PID:g1000
A;Experimental source: strain Wistar; tissue brain
C;Genetics:
A;Gene: limk-2a
C;Superfamily: LIM protein kinase; LIM metal-binding repeat homology; protein kinase ho
C;Keywords: alternative splicing; ATP; duplication; phosphotransferase; serine/threonin
F;12-63/Domain: LIM metal-binding repeat homology <LIM1>
F;72-124/Domain: LIM metal-binding repeat homology <LIM2>
F;329-608/Domain: protein kinase homology <KIN>
F;337-345/Region: protein kinase ATP-binding motif

Query Match 73.3%; Score 33; DB 1; Length 638;
Best Local Similarity 55.6%; Pred. No. 85;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IMDQVPFSV 9
: ||:|:||
Db 142 VQDQLPYSV 150

RESULT 11
JC5813
LIM motif-containing protein kinase (EC 2.7.1.-) 2, splice form a - mouse
C;Species: Mus musculus (house mouse)
C;Date: 04-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C;Accession: JC5813
R;Koshimizu, U.; Takahashi, H.; Yoshida, M.C.; Nakamura, T.
Biochem. Biophys. Res. Commun. 241, 243-250, 1997
A;Title: cDNA cloning, genomic organization, and chromosomal localization of the mouse
A;Reference number: JC5813; MUID:98086337; PMID:9425257
A;Accession: JC5813
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-638 <KOS>
A;Cross-references: UNIPROT:O54785; DDBJ:AB008117; NID:g2789463; PID:g2789464
A;Experimental source: embryo
C;Comment: This enzyme is involved in a signal transduction pathway to regulate various
C;Genetics:
A;Gene: Limk2a
A;Map position: 1d
C;Superfamily: LIM protein kinase; LIM metal-binding repeat homology; protein kinase ho
C;Keywords: alternative splicing; phosphotransferase
F;12-63/Domain: LIM metal-binding repeat homology <LIM1>
F;72-124/Domain: LIM metal-binding repeat homology <LIM2>
F;329-608/Domain: protein kinase homology <KIN>
F;338-601/Domain: protein kinase #status predicted <PKD>

Query Match 73.3%; Score 33; DB 2; Length 638;
Best Local Similarity 55.6%; Pred. No. 85;

```
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              : ||:|:||
Db        142 VQDQLPYSV 150


RESULT 12
E46178
probable transcription factor fork head domain 5 (FD5) - fruit fly (Drosophila melanogas
C;Species: Drosophila melanogaster
C;Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 16-Aug-2004
C;Accession: E46178
R;Hacker, U.; Grossniklaus, U.; Gehring, W.J.; Jackle, H.
Proc. Natl. Acad. Sci. U.S.A. 89, 8754-8758, 1992
A;Title: Developmentally regulated Drosophila gene family encoding the fork head domain.
A;Reference number: A46178; MUID:92409595; PMID:1356269
A;Accession: E46178
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-128 <HAC>
A;Cross-references: UNIPROT:P32029; GB:M96444; NID:g157431; PIDN:AAF02178.1; PID:g604218
A;Note: sequence extracted from NCBI backbone (NCBIP:114227)
C;Superfamily: fork head DNA-binding domain homology
F;16-107/Domain: fork head DNA-binding domain homology <FHD>

  Query Match             71.1%;  Score 32;  DB 2;  Length 128;
  Best Local Similarity   85.7%;  Pred. No. 24;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPF 7
              |||| ||
Db         46 IMDQFPF 52


RESULT 13
S65051
low molecular weight heat shock protein precursor (clone Hsp22.5), endoplasmic reticulum
C;Species: Glycine max (soybean)
C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S65051
R;LaFayette, P.R.; Nagao, R.T.; O'Grady, K.; Vierling, E.; Key, J.L.
Plant Mol. Biol. 30, 159-169, 1996
A;Title: Molecular characterization of cDNAs encoding low-molecular-weight heat shock pr
A;Reference number: S65049; MUID:96197406; PMID:8616233
A;Accession: S65051
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-197 <LAF>
A;Cross-references: UNIPROT:Q39820; EMBL:U21724
C;Keywords: heat shock

  Query Match             71.1%;  Score 32;  DB 2;  Length 197;
  Best Local Similarity   44.4%;  Pred. No. 38;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              :::|:|| |
Db         51 VLEQIPFGV 59


RESULT 14
S72398
low molecular weight heat shock protein precursor (clone Hsp22.5), endoplasmic reticulum
C;Species: Glycine max (soybean)
C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S72398
R;LaFayette, P.R.
submitted to the EMBL Data Library, February 1995
A;Reference number: S72398
A;Accession: S72398
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-197 <LAF>
A;Cross-references: UNIPROT:Q39820; EMBL:U21724; NID:g710435; PIDN:AAB03098.1; PID:g710

  Query Match             71.1%;  Score 32;  DB 2;  Length 197;
  Best Local Similarity   44.4%;  Pred. No. 38;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              :::|:|| |
Db         51 VLEQIPFGV 59


RESULT 15
C90433
ABC transporter, ATP binding protein SSO2600 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 16-Aug-2004
C;Accession: C90433
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: C90433
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-249 <KUR>
A;Cross-references: UNIPROT:Q97VM2; GB:AE006641; NID:g13815904; PIDN:AAK42722.1; GSPDB:
C;Genetics:
A;Gene: SSO2600
C;Superfamily: ATP-binding cassette homology

  Query Match             71.1%;  Score 32;  DB 2;  Length 249;
  Best Local Similarity   85.7%;  Pred. No. 49;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPF 7
              ||| |||
Db         54 IMDMVPF 60


Search completed: December  8, 2004, 17:13:37
Job time : 5.7561 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2004, 17:06:18 ; Search time 6.91463 Seconds
(without alignments)
748.900 Million cell updates/sec

Title: US-10-073-301A-2
Perfect score: 45
Sequence: 1 IMDQVPFSV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	88.9	199	2	Q8IG08	Q8ig08 caenorhabdi
2	40	88.9	283	2	Q21305	Q21305 caenorhabdi
3	39	86.7	461	2	O97884	O97884 equus cabal
4	39	86.7	491	1	PM17_BOVIN	Q06154 bos taurus
5	39	86.7	626	1	PM17_MOUSE	Q60696 mus musculu
6	39	86.7	626	2	Q9CZB2	Q9czb2 mus musculu
7	39	86.7	661	1	PM17_HUMAN	P40967 homo sapien
8	39	86.7	661	2	AAP35866	Aap35866 homo sapi
9	37	82.2	684	2	Q7ZVU9	Q7zvu9 brachydanio
10	36	80.0	964	2	Q72BS2	Q72bs2 desulfovibr
11	36	80.0	964	2	AAS96040	Aas96040 desulfovi
12	35	77.8	316	2	Q7XUY6	Q7xuy6 oryza sativ
13	35	77.8	417	2	Q9F6E7	Q9f6e7 streptomyce
14	35	77.8	638	2	Q7MJG0	Q7mjg0 vibrio vuln
15	34	75.6	149	2	Q94CH0	Q94ch0 zea mays (m
16	34	75.6	342	2	Q6PBN3	Q6pbn3 brachydanio
17	34	75.6	342	2	AAH59645	Aah59645 brachydan
18	34	75.6	345	1	RDS1_XENLA	O42581 xenopus lae
19	34	75.6	345	2	Q7SY68	Q7sy68 xenopus lae
20	34	75.6	402	2	Q7MD26	Q7md26 vibrio vuln
21	34	75.6	402	2	Q8D607	Q8d607 vibrio vuln
22	34	75.6	3167	2	Q7QQ87	Q7qq87 giardia lam
23	33	73.3	98	2	Q9QWM3	Q9qwm3 mus musculu
24	33	73.3	239	2	Q9ZQW1	Q9zqw1 triticum ae
25	33	73.3	315	2	Q8IRR3	Q8irr3 drosophila
26	33	73.3	337	2	Q7LYI2	Q7lyi2 pyrococcus
27	33	73.3	337	2	Q9HH23	Q9hh23 thermococcu
28	33	73.3	337	2	Q9HH28	Q9hh28 pyrococcus
29	33	73.3	390	2	Q82DI8	Q82di8 streptomyce
30	33	73.3	422	2	Q6NX08	Q6nx08 brachydanio
31	33	73.3	422	2	Q803T7	Q803t7 brachydanio
32	33	73.3	422	2	AAH67331	Aah67331 brachydan
33	33	73.3	439	2	Q9PMK6	Q9pmk6 campylobact
34	33	73.3	556	2	Q98KS1	Q98ks1 rhizobium l
35	33	73.3	614	2	Q7UTZ4	Q7utz4 rhodopirell
36	33	73.3	617	2	AAH65578	Aah65578 rattus no
37	33	73.3	638	1	LIK2_MOUSE	O54785 mus musculu
38	33	73.3	638	1	LIK2_RAT	P53670 rattus norv
39	33	73.3	669	2	Q86ZP6	Q86zp6 alternaria
40	33	73.3	1145	2	Q7Z8T3	Q7z8t3 cochliobolu
41	33	73.3	1239	2	Q6SLC7	Q6slc7 cochliobolu
42	33	73.3	1239	2	AAR29899	Aar29899 cochliobo
43	33	73.3	1329	2	Q7Z8T6	Q7z8t6 cochliobolu
44	33	73.3	1329	2	Q7Z8T7	Q7z8t7 cochliobolu
45	33	73.3	1329	2	Q7Z8T8	Q7z8t8 cochliobolu

ALIGNMENTS

RESULT 1
Q8IG08
ID Q8IG08 PRELIMINARY; PRT; 199 AA.
AC Q8IG08;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Sox (Mammalian sry box) family protein 2, isoform b.
GN Name=sox-2; ORFNames=K08A8.2;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Pauley A.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U38377; AAN72421.1; -.
DR HSSP; P48432; 1GT0.
DR WormPep; K08A8.2b; CE31200.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009071; HMG-box.
DR InterPro; IPR000910; HMG_12_box.
DR Pfam; PF00505; HMG_box; 1.
DR SMART; SM00398; HMG; 1.
DR PROSITE; PS50118; HMG_BOX_2; 1.
SQ SEQUENCE 199 AA; 22734 MW; D737C6B759ED8944 CRC64;

Query Match 88.9%; Score 40; DB 2; Length 199;
Best Local Similarity 66.7%; Pred. No. 2.5;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IMDQVPFSV 9
:|||:|||:

```
Db        101 VMDQIPFSL 109


RESULT 2
Q21305
ID   Q21305      PRELIMINARY;      PRT;   283 AA.
AC   Q21305;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Sox (Mammalian sry box) family protein 2, isoform a.
GN   Name=sox-2; ORFNames=K08A8.2;
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   Wilson R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Pauley A.;
RL   Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Wilson R.;
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U38377; AAA79747.2; -.
DR   PIR; T16581; T16581.
DR   HSSP; P48432; 1GT0.
DR   WormPep; K08A8.2a; CE28595.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR009071; HMG-box.
DR   InterPro; IPR000910; HMG_12_box.
DR   Pfam; PF00505; HMG_box; 1.
DR   SMART; SM00398; HMG; 1.
DR   PROSITE; PS50118; HMG_BOX_2; 1.
SQ   SEQUENCE   283 AA;  32169 MW;  71F46FD3BDDAAFA4 CRC64;

  Query Match             88.9%;  Score 40;  DB 2;  Length 283;
  Best Local Similarity   66.7%;  Pred. No. 3.6;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              :|||:|||:
Db        185 VMDQIPFSL 193


RESULT 3
O97884
ID   O97884      PRELIMINARY;      PRT;   461 AA.
AC   O97884;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Melanocyte protein 17 (Fragment).
GN   Name=PMEL17;
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RA   Rieder S., Stricker C., Joerg H., Dummer R., Stranzinger G.;
RT   "A comparative genetic approach for the investigation of ageing grey
RT   house melanoma.";
RL   J. Anim. Breed. Genet. 117:73-82(2000).
DR   EMBL; AF076780; AAC97108.1; -.
DR   InterPro; IPR000601; PKD.
DR   Pfam; PF00801; PKD; 1.
DR   SMART; SM00089; PKD; 1.
DR   PROSITE; PS50093; PKD; 1.
FT   NON_TER       1      1
FT   NON_TER     461    461
SQ   SEQUENCE   461 AA;  49334 MW;  12752AF6C1EC373D CRC64;

  Query Match             86.7%;  Score 39;  DB 2;  Length 461;
  Best Local Similarity   88.9%;  Pred. No. 10;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              | |||||||
Db         18 ITDQVPFSV 26


RESULT 4
PM17_BOVIN
ID   PM17_BOVIN     STANDARD;      PRT;   491 AA.
AC   Q06154;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Melanocyte protein Pmel 17 (Retinal pigment epithelial-specific
DE   protein) (Fragment).
GN   Name=SILV; Synonyms=PMEL17, RPE1;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Retina;
RX   MEDLINE=93122163; PubMed=1478275;
RA   Kim R.Y., Wistow G.J.;
RT   "The cDNA RPE1 and monoclonal antibody HMB-50 define gene products
RT   preferentially expressed in retinal pigment epithelium.";
RL   Exp. Eye Res. 55:657-662(1992).
CC   -!- FUNCTION: Could be a melanogenic enzyme (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC   -!- TISSUE SPECIFICITY: Retinal pigment epithelium.
CC   -!- SIMILARITY: Belongs to the Pmel-17/NMB family.
CC   -!- SIMILARITY: Contains 1 PKD domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M81193; AAA30419.1; -.
DR   PIR; A49179; A49179.
DR   InterPro; IPR000601; PKD.
DR   Pfam; PF00801; PKD; 1.
DR   SMART; SM00089; PKD; 1.
DR   PROSITE; PS50093; PKD; 1.
KW   Glycoprotein; Melanin biosynthesis; Repeat; Transmembrane.
FT   NON_TER       1      1
FT   DOMAIN       <1    423       Extracellular (Potential).
FT   TRANSMEM    424    444       Potential.
```

```
FT   DOMAIN       445    491       Cytoplasmic (Potential).
FT   DOMAIN        60    150       PKD.
FT   DOMAIN       148    256       8 X 13 AA approximate tandem repeats.
FT   REPEAT       148    160       1.
FT   REPEAT       161    173       2.
FT   REPEAT       174    186       3.
FT   REPEAT       187    199       4.
FT   REPEAT       200    212       5.
FT   REPEAT       213    225       6.
FT   REPEAT       232    243       7.
FT   REPEAT       244    256       8.
FT   DOMAIN       304    394       Cys-rich.
FT   CARBOHYD     269    269       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     396    396       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   491 AA;  51669 MW;  2BBFE5DFBD397D6D CRC64;

  Query Match             86.7%;  Score 39;  DB 1;  Length 491;
  Best Local Similarity   88.9%;  Pred. No. 11;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              | |||||||
Db         52 ITDQVPFSV 60


RESULT 5
PM17_MOUSE
ID   PM17_MOUSE     STANDARD;      PRT;   626 AA.
AC   Q60696;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Melanocyte protein Pmel 17 precursor (Silver locus protein).
GN   Name=Silv; Synonyms=Pmel17, D10H12S53E, Si;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND VARIANTS SILVER.
RC   STRAIN=C57BL/6; TISSUE=Skin;
RX   MEDLINE=95175358; PubMed=7870580;
RA   Kwon B.S., Halaban R., Ponnazhagan S., Kim K., Chintamaneni C.,
RA   Bennett D., Pickard R.T.;
RT   "Mouse silver mutation is caused by a single base insertion in the
RT   putative cytoplasmic domain of Pmel 17.";
RL   Nucleic Acids Res. 23:154-158(1995).
CC   -!- FUNCTION: Could be a melanogenic enzyme.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC   -!- TISSUE SPECIFICITY: Preferentially expressed in melanocytes.
CC   -!- DISEASE: Defects in Silv are the cause of the silver coat color
CC       which seems to be due to premature death of pigment cells during
CC       the hair cycle.
CC   -!- SIMILARITY: Belongs to the Pmel-17/NMB family.
CC   -!- SIMILARITY: Contains 1 PKD domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U14133; AAA69538.1; -.
DR   PIR; S53871; S53871.
DR   MGD; MGI:98301; Si.
DR   InterPro; IPR000601; PKD.
DR   Pfam; PF00801; PKD; 1.
DR   PROSITE; PS50093; PKD; 1.
KW   Disease mutation; Glycoprotein; Melanin biosynthesis; Repeat; Signal;
KW   Transmembrane.
FT   SIGNAL        1     24       Potential.
FT   CHAIN        25    626       Melanocyte protein Pmel 17.
FT   DOMAIN       25    562       Extracellular (Potential).
FT   TRANSMEM    563    583       Potential.
FT   DOMAIN       584    626       Cytoplasmic (Potential).
FT   DOMAIN       255    292       PKD.
FT   DOMAIN       315    411       7 X 13 AA APPROXIMATE TANDEM REPEATS.
FT   REPEAT       315    327       1.
FT   REPEAT       328    340       2.
FT   REPEAT       341    353       3.
FT   REPEAT       354    366       4.
FT   REPEAT       367    379       5.
FT   REPEAT       380    392       6.
FT   REPEAT       393    411       7.
FT   CARBOHYD      81     81       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     106    106       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     111    111       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     535    535       N-linked (GlcNAc. . .) (Potential).
FT   VARIANT      170    170       S -> L (in silver).
FT   VARIANT      175    175       R -> G (in silver).
FT   VARIANT      373    373       D -> N (in silver).
FT   VARIANT      471    471       F -> S (in silver).
FT   VARIANT      603    626       AAPASGLRARGLGENSPLLSGQQV -> SSASLRSSRPRPW
FT                                 RKQPAPQWTAGLIILKAPWISWG (in silver).
SQ   SEQUENCE   626 AA;  65980 MW;  7AB941D2E3FB1044 CRC64;

  Query Match             86.7%;  Score 39;  DB 1;  Length 626;
  Best Local Similarity   88.9%;  Pred. No. 14;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              | |||||||
Db        209 ITDQVPFSV 217


RESULT 6
Q9CZB2
ID   Q9CZB2       PRELIMINARY;      PRT;   626 AA.
AC   Q9CZB2;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
DE   enriched library, clone:2810025C24 product:silver, full insert
DE   sequence.
GN   Name=Si;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=99279253; PubMed=10349636;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=21085660; PubMed=11217851;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=20499374; PubMed=11042159;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=20530913; PubMed=11076861;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RA   Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA   Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA   Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA   Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA   Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA   Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA   Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA   Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA   Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA   Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA   Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK012808; BAB28486.1; -.
DR   MGD; MGI:98301; Si.
DR   InterPro; IPR000601; PKD.
DR   Pfam; PF00801; PKD; 1.
DR   SMART; SM00089; PKD; 1.
DR   PROSITE; PS50093; PKD; 1.
SQ   SEQUENCE   626 AA;  66301 MW;  7EC0A06C63212674 CRC64;

  Query Match             86.7%;  Score 39;  DB 2;  Length 626;
  Best Local Similarity   88.9%;  Pred. No. 14;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              | ||||||||
Db        208 ITDQVPFSV 216


RESULT 7
PM17_HUMAN
ID   PM17_HUMAN     STANDARD;      PRT;   661 AA.
AC   P40967; Q12763; Q14448; Q14817; Q16565;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Melanocyte protein Pmel 17 precursor (Melanocyte lineage-specific
DE   antigen GP100) (Melanoma-associated ME20 antigen) (ME20M/ME20S) (ME20-
DE   M/ME20-S) (95 kDa melanocyte-specific secreted glycoprotein).
GN   Name=SILV; Synonyms=PMEL17, D12S53E;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92021023; PubMed=1924386;
RA   Kwon B.S., Chintamaneni C., Kozak C.A., Copeland N.G., Gilbert D.J.,
RA   Jenkins N.A., Barton D., Prancke U., Kobayashi Y., Kim K.-K.;
RT   "A melanocyte-specific gene, Pmel 17, maps near the silver coat color
RT   locus on mouse chromosome 10 and is in a syntenic region on human
RT   chromosome 12.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:9228-9232(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94327568; PubMed=7519602;
RA   Adema G.J., de Boer A.J., Vogel A.M., Loenen W.A., Figdor C.G.;
RT   "Molecular characterization of the melanocyte lineage-specific antigen
RT   gp100.";
RL   J. Biol. Chem. 269:20126-20133(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96154052; PubMed=8592076;
RA   Bailin T., Lee S.-T., Spritz R.A.;
RT   "Genomic organization and sequence of D12S53E (Pmel 17), the human
RT   homologue of the mouse silver (si) locus.";
RL   J. Invest. Dermatol. 106:24-27(1996).
RN   [4]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 25-53.
RX   MEDLINE=94235165; PubMed=8179825;
RA   Maresh G.A., Marken J.S., Neubauer M., Aruffo A., Hellstroem I.,
RA   Hellstroem K.E., Marquardt H.;
RT   "Cloning and expression of the gene for the melanoma-associated ME20
RT   antigen.";
RL   DNA Cell Biol. 13:87-95(1994).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96314705; PubMed=8739560;
RA   Kim K.K., Youn B.S., Heng H.H., Shi X.-M., Tsui L.-C., Lee Z.H.,
RA   Pickard R.T., Kwon B.S.;
RT   "Genomic organization and FISH mapping of human Pmel 17, the putative
RT   silver locus.";
RL   Pigment Cell Res. 9:42-48(1996).
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Vogel A.;
RL   Submitted (NOV-1990) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Could be a melanogenic enzyme. Could represent an
CC       oncofetal self-antigen that is normally expressed at low levels in
CC       quiescent adult melanocytes but overexpressed by proliferating
CC       neonatal melanocytes and during tumor growth. Release of the
CC       soluble form, ME20-S, could protect tumor cells from antibody
CC       mediated immunity.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential). There
CC       is also a secreted soluble form, ME20-S, probably product of
CC       proteolytic cleavage.
CC   -!- TISSUE SPECIFICITY: Preferentially expressed in melanomas. Some
CC       expression was found in dysplastic nevi. Not found in normal
```

```
CC       tissues nor in carcinomas.
CC   -!- SIMILARITY: Belongs to the Pmel-17/NMB family.
CC   -!- SIMILARITY: Contains 1 PKD domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M77348; AAA60121.1; -.
DR   EMBL; S73003; AAC60634.1; -.
DR   EMBL; U31799; AAB00386.1; -.
DR   EMBL; U31808; AAB00386.1; JOINED.
DR   EMBL; U31807; AAB00386.1; JOINED.
DR   EMBL; U31797; AAB00386.1; JOINED.
DR   EMBL; U31798; AAB00386.1; JOINED.
DR   EMBL; U01874; AAA18479.1; -.
DR   EMBL; U20093; AAB19181.1; -.
DR   EMBL; U19491; AAB19181.1; JOINED.
DR   EMBL; M32295; AAA35930.1; ALT_INIT.
DR   EMBL; BC001414; AAH01414.1; -.
DR   PIR; I38400; I38400.
DR   Genew; HGNC:10880; SILV.
DR   MIM; 155550; -.
DR   GO; GO:0005886; C:plasma membrane; TAS.
DR   InterPro; IPR000601; PKD.
DR   Pfam; PF00801; PKD; 1.
DR   PROSITE; PS50093; PKD; 1.
KW   Antigen; Direct protein sequencing; Glycoprotein;
KW   Melanin biosynthesis; Repeat; Signal; Transmembrane.
FT   SIGNAL        1     24
FT   CHAIN        25    661       Melanocyte protein Pmel 17.
FT   DOMAIN       25    595       Extracellular (Potential).
FT   TRANSMEM    596    616       Potential.
FT   DOMAIN      617    661       Cytoplasmic (Potential).
FT   DOMAIN      255    292       PKD.
FT   DOMAIN      315    444       10 X 13 AA approximate tandem repeats.
FT   REPEAT      315    327       1.
FT   REPEAT      328    340       2.
FT   REPEAT      341    353       3.
FT   REPEAT      354    366       4.
FT   REPEAT      367    379       5.
FT   REPEAT      380    392       6.
FT   REPEAT      393    405       7.
FT   REPEAT      406    418       8.
FT   REPEAT      419    431       9.
FT   REPEAT      432    444       10.
FT   CARBOHYD     81     81       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    106    106       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    111    111       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    321    321       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    568    568       N-linked (GlcNAc. . .) (Potential).
FT   CONFLICT    274    274       L -> P (in Ref. 1 and 5).
FT   CONFLICT    587    587       P -> PVPGILLT (in Ref. 1).
FT   CONFLICT    592    592       G -> GG (in Ref. 4).
FT   CONFLICT    597    597       P -> R (in Ref. 1).
FT   CONFLICT    642    661       RIPCSCPIGENSPLLSGQQV -> ASSALVPLVRIAPSSVG
FT                                SRSEYSHMML (in Ref. 6).
SQ   SEQUENCE   661 AA;  70255 MW;  8A904FAB16715653 CRC64;

  Query Match             86.7%;  Score 39;  DB 1;  Length 661;
  Best Local Similarity   88.9%;  Pred. No. 15;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              | |||||||
Db        209 ITDQVPFSV 217


RESULT 8
```

```
AAP35866
ID   AAP35866    PRELIMINARY;      PRT;   661 AA.
AC   AAP35866;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   Silver homolog (Mouse).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA   Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA   Phelan M., Farmer A.;
RT   "Cloning of human full-length CDSs in BD Creator(TM) System Donor
RT   vector.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BT007202; AAP35866.1; -.
SQ   SEQUENCE   661 AA;  70255 MW;  8A904FAB16715653 CRC64;

  Query Match             86.7%;  Score 39;  DB 2;  Length 661;
  Best Local Similarity   88.9%;  Pred. No. 15;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              | |||||||
Db        209 ITDQVPFSV 217


RESULT 9
Q7ZVU9
ID   Q7ZVU9      PRELIMINARY;      PRT;   684 AA.
AC   Q7ZVU9;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Nnp1 protein (Fragment).
GN   Name=nnp1;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
RA   Strausberg R.;
```

```
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC045402; AAH45402.1; -.
DR   ZFIN; ZDB-GENE-030131-9837; nnp1.
DR   InterPro; IPR010301; Nop52.
DR   Pfam; PF05997; Nop52; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   684 AA;  77380 MW;  DC5014B08FECBC0C CRC64;

  Query Match             82.2%;  Score 37;  DB 2;  Length 684;
  Best Local Similarity   66.7%;  Pred. No. 44;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              |:|||||::
Db        232 IVDQVPFAI 240


RESULT 10
Q72BS2
ID   Q72BS2      PRELIMINARY;      PRT;   964 AA.
AC   Q72BS2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   HAMP domain protein.
GN   OrderedLocusNames=DVU1562;
OS   Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS   8303).
OC   Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC   Desulfovibrionaceae; Desulfovibrio.
OX   NCBI_TaxID=882;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=15077118; DOI=10.1038/nbt959;
RA   Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA   Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA   Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA   Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA   Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D.,
RA   Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA   Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT   "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT   Desulfovibrio vulgaris Hildenborough.";
RL   Nat. Biotechnol. 22:554-559(2004).
DR   EMBL; AE017314; AAS96040.1; -.
DR   TIGR; DVU1562; -.
DR   InterPro; IPR003018; GAF.
DR   InterPro; IPR003660; HAMP.
DR   InterPro; IPR003607; Met_phos_hydro.
DR   Pfam; PF01590; GAF; 1.
DR   Pfam; PF00672; HAMP; 1.
DR   SMART; SM00471; HDc; 1.
KW   Complete proteome.
SQ   SEQUENCE   964 AA;  107427 MW;  C3E9682BBE43CC79 CRC64;

  Query Match             80.0%;  Score 36;  DB 2;  Length 964;
  Best Local Similarity   75.0%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFS 8
              ::||||||
Db        542 VIDQVPFS 549


RESULT 11
AAS96040
ID   AAS96040    PRELIMINARY;      PRT;   964 AA.
AC   AAS96040;
DT   27-APR-2004 (TrEMBLrel. 27, Created)
DT   27-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT   11-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE   HAMP domain protein.
GN   DVU1562.
OS   Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS   8303).
OC   Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC   Desulfovibrionaceae; Desulfovibrio.
OX   NCBI_TaxID=882;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=15077118;
RA   Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA   Kolonay J.F., Eisen J.A., Ward N., Methe B.A., Brinkac L.M.,
RA   Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA   Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA   Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D.,
RA   Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA   Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT   "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT   Desulfovibrio vulgaris Hildenborough.";
RL   Nat. Biotechnol. 22:554-559(2004).
DR   EMBL; AE017314; AAS96040.1; -.
DR   TIGR; DVU1562; -.
SQ   SEQUENCE   964 AA;  107427 MW;  C3E9682BBE43CC79 CRC64;

  Query Match             80.0%;  Score 36;  DB 2;  Length 964;
  Best Local Similarity   75.0%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFS 8
              ::||||||
Db        542 VIDQVPFS 549


RESULT 12
Q7XUY6
ID   Q7XUY6      PRELIMINARY;      PRT;   316 AA.
AC   Q7XUY6;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   OSJNBb0048E02.11 protein.
GN   Name=OSJNBb0048E02.11;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=12447439;
RA   Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA   Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA   Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA   Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA   Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA   Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA   Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA   Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA   Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA   Han B.;
RT   "Sequence and analysis of rice chromosome 4.";
RL   Nature 420:316-320(2002).
DR   EMBL; AL606653; CAD40935.1; -.
DR   Gramene; Q7XUY6; -.
DR   GO; GO:0008080; F:N-acetyltransferase activity; IEA.
DR   InterPro; IPR000182; GCN5acetyl_trans.
DR   Pfam; PF00583; Acetyltransf_1; 1.
SQ   SEQUENCE   316 AA;  34542 MW;  36C799C4C038EC66 CRC64;

  Query Match             77.8%;  Score 35;  DB 2;  Length 316;
  Best Local Similarity   66.7%;  Pred. No. 53;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              |:|||||::
```

```
Db        256 IIDQVPFNM 264


RESULT 13
Q9F6E7
ID   Q9F6E7      PRELIMINARY;      PRT;   417 AA.
AC   Q9F6E7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   PKSA beta-ketoacylsynthase subunit beta.
OS   Streptomyces collinus.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=42684;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM2012;
RA   Martin R., Bailey J.E., Minas W.;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the beta-ketoacyl-ACP synthases family.
DR   EMBL; AF293354; AAG26880.1; -.
DR   HSSP; P73283; 1E5M.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR   InterPro; IPR000794; Ketoacyl_synth.
DR   Pfam; PF00109; ketoacyl-synt; 1.
DR   Pfam; PF02801; Ketoacyl-synt_C; 1.
KW   Transferase.
SQ   SEQUENCE   417 AA;  43370 MW;  483F4BE4A31205AD CRC64;

  Query Match             77.8%;  Score 35;  DB 2;  Length 417;
  Best Local Similarity   66.7%;  Pred. No. 71;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              : |:|||||
Db         98 VTDEVPFSV 106


RESULT 14
Q7MJG0
ID   Q7MJG0      PRELIMINARY;      PRT;   638 AA.
AC   Q7MJG0;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Type I restriction-modification system methyltransferase subunit.
GN   Name=VV2202;
OS   Vibrio vulnificus (strain YJ016).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=196600;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=14656965;
RA   Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA   Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA   Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT   "Comparative genome analysis of Vibrio vulnificus, a marine
RT   pathogen.";
RL   Genome Res. 13:2577-2587(2003).
DR   EMBL; AP005338; BAC94966.1; -.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR   GO; GO:0009007; F:site-specific DNA-methyltransferase (adenin. . .; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006306; P:DNA methylation; IEA.
DR   GO; GO:0006304; P:DNA modification; IEA.
DR   InterPro; IPR003665; Methylase_M.
DR   InterPro; IPR002296; N12N6_mtfrase.
DR   InterPro; IPR003356; N6_DNA_Mtase.
DR   InterPro; IPR000055; Rest_mod_DNA.
DR   Pfam; PF02506; Methylase_M; 1.
DR   Pfam; PF01420; Methylase_S; 1.
DR   Pfam; PF02384; N6_Mtase; 1.
DR   PRINTS; PR00507; N12N6MTFRASE.
KW   Methyltransferase; Transferase.
SQ   SEQUENCE   638 AA;  71784 MW;  4D93E4496B30DF66 CRC64;

  Query Match             77.8%;  Score 35;  DB 2;  Length 638;
  Best Local Similarity   55.6%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              |:||:|| :
Db        110 ILDQIPFRI 118


RESULT 15
Q94CH0
ID   Q94CH0      PRELIMINARY;      PRT;   149 AA.
AC   Q94CH0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Seven transmembrane protein Mlo5 (Fragment).
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22456933; PubMed=12569425;
RA   Devoto A., Hartmann H.A., Piffanelli P., Elliott C., Simmons C.,
RA   Taramino G., Goh C.-S., Cohen F.E., Emerson B.C., Schulze-Lefert P.,
RA   Panstruga R.;
RT   "Molecular phylogeny and evolution of the plant-specific seven-
RT   transmembrane MLO family.";
RL   J. Mol. Evol. 56:77-88(2003).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Briggs S.P., Simmons C.R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY029316; AAK38341.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0008219; P:cell death; IEA.
DR   InterPro; IPR004326; Mlo.
DR   Pfam; PF03094; Mlo; 1.
KW   Transmembrane.
FT   NON_TER       1       1
SQ   SEQUENCE   149 AA;  16260 MW;  D037057B000EE8A4 CRC64;

  Query Match             75.6%;  Score 34;  DB 2;  Length 149;
  Best Local Similarity   77.8%;  Pred. No. 39;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 IMDQVPFSV 9
              ||| ||| |
Db         28 IMDHVPFIV 36


Search completed: December  8, 2004, 17:09:51
Job time : 13.9146 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2004, 17:06:18 ; Search time 182.085 Seconds
(without alignments)
748.900 Million cell updates/sec

Title: US-10-073-301A-9
Perfect score: 1272
Sequence: 1 QVKLQESGGGLVKPGGSLKL.....TYYCQEWSGYPYTFGGGTKL 237

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	936.5	73.6	255	2	Q6KB05	Q6kb05 mus musculu
2	936.5	73.6	255	2	CAG34081	Cag34081 mus muscu
3	838	65.9	298	2	Q9QYF0	Q9qyf0 synthetic c
4	710.5	55.9	241	2	Q921A6	Q921a6 mus musculu
5	692.5	54.4	243	2	Q7TQM2	Q7tqm2 mus musculu
6	588	46.2	218	2	Q925S1	Q925s1 mus musculu
7	543.5	42.7	487	2	Q99KA4	Q99ka4 mus musculu
8	531	41.7	119	2	Q920E7	Q920e7 mus musculu
9	502	39.5	486	2	Q91Z07	Q91z07 mus musculu
10	496	39.0	480	2	Q91XE1	Q91xe1 mus musculu
11	487	38.3	471	2	AAH24289	Aah24289 homo sapi
12	487	38.3	479	2	Q91WP5	Q91wp5 mus musculu
13	478	37.6	108	2	AAR11048	Aar11048 mus muscu
14	477.5	37.5	606	2	Q6GMY2	Q6gmy2 homo sapien
15	475	37.3	464	2	Q6MZU6	Q6mzu6 homo sapien
16	475	37.3	464	2	CAE45931	Cae45931 homo sapi
17	474.5	37.3	124	2	BAD00233	Bad00233 camelus d
18	473	37.2	473	2	Q91Z05	Q91z05 mus musculu
19	471.5	37.1	478	2	Q6PI81	Q6pi81 homo sapien
20	471.5	37.1	478	2	AAH41037	Aah41037 homo sapi
21	470	36.9	106	2	Q9U410	Q9u410 schistosoma
22	470	36.9	124	2	BAD00534	Bad00534 camelus d
23	469.5	36.9	120	2	BAD00231	Bad00231 camelus d
24	469.5	36.9	126	2	BAD00225	Bad00225 camelus d
25	468.5	36.8	125	2	BAD00491	Bad00491 camelus d
26	468	36.8	122	2	BAD00446	Bad00446 camelus d
27	467	36.7	98	1	HV57_MOUSE	P18528 mus musculu
28	467	36.7	119	2	AAL35865	Aal35865 lama glam
29	467	36.7	126	2	BAD00420	Bad00420 camelus d
30	467	36.7	126	2	BAD00529	Bad00529 camelus d
31	466	36.6	128	2	BAD00406	Bad00406 camelus d
32	465.5	36.6	470	2	Q6PJA4	Q6pja4 homo sapien
33	465.5	36.6	470	2	AAH18747	Aah18747 homo sapi
34	465	36.6	128	2	BAD00444	Bad00444 camelus d
35	464.5	36.5	119	2	BAD00399	Bad00399 camelus d
36	464	36.5	122	2	BAD00549	Bad00549 camelus d
37	464	36.5	126	2	BAD00510	Bad00510 camelus d
38	463	36.4	117	2	AAL35875	Aal35875 lama glam
39	462.5	36.4	136	1	HV16_MOUSE	P01783 mus musculu
40	462.5	36.4	613	2	Q8WUK1	Q8wuk1 homo sapien
41	460	36.2	117	1	HV55_MOUSE	P18526 mus musculu
42	460	36.2	123	2	BAD00234	Bad00234 camelus d
43	459.5	36.1	97	1	HV56_MOUSE	P18527 mus musculu
44	459	36.1	437	2	Q9R1A4	Q9r1a4 mus musculu
45	458.5	36.0	125	2	BAD00448	Bad00448 camelus d

ALIGNMENTS

RESULT 1
Q6KB05
ID Q6KB05 PRELIMINARY; PRT; 255 AA.
AC Q6KB05;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ScFv B8E5 protein (Fragment).
GN Name=scFv B8E5;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/c;
RA Peter J.C., Wallukat G., Tugler J., Maurice D., Roegel J.C.,
RA Briand J.P., Hoebeke J.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ746180; CAG34081.1; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 2.
DR SMART; SM00409; IG; 2.
DR SMART; SM00406; IGv; 2.
DR PROSITE; PS50835; IG_LIKE; 2.
FT NON_TER 1 1
SQ SEQUENCE 255 AA; 27445 MW; B68BD38395DF713B CRC64;

Query Match 73.6%; Score 936.5; DB 2; Length 255;
Best Local Similarity 72.9%; Pred. No. 1.6e-62;
Matches 180; Conservative 23; Mismatches 33; Indels 11; Gaps 3;

```
Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              ||:||:||| |||||||||:|||||||||||||||||||||||||||||:||||||||
Db          1 QVQLQQSGGDLVKPGGSLKVSCAASGFTFSSYGMSWVRQTPDKRLEWVATITSGGSYTYY 60

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARG---NWEGWYFDVWGQGTTVTV 117
              ||||||||||||||||||||||||||||||||||||||     ::|  || ||||||:||
Db         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARHINYRYDG-AFDVWGQGTTLTV 119

Qy        118 SSGGGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSI-------RYIYWYQ 170
              |||||||||||||||:| : |||: :| | ||:| |:| :| |:         |: |||
Db        120 SSGGGGSGGGGSGGGGSDIVMAQSPSSLSVSAGEKVIMSCKSSQSLLNSRNQKNYLAWYQ 179

Qy        171 QKPGSSPRLLIYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYT 230
              |||| ||:|||| |    ||| ||:|||||| ::|||: ::||| | ||||    || |
Db        180 QKPGQSPKLLIYGASTRESGVPDRFTGSGSGTDFTLTISSVQAEDLAVYYCQNDHSYPLT 239

Qy        231 FGGGTKL 237
              || ||||
Db        240 FGAGTKL 246
```

```
RESULT 2
CAG34081
ID   CAG34081    PRELIMINARY;      PRT;   255 AA.
AC   CAG34081;
DT   01-JUN-2004 (TrEMBLrel. 27, Created)
DT   01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT   01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE   ScFv B8E5 protein (Fragment).
GN   SCFV B8E5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=balb/c;
RA   Peter J.C., Wallukat G., Tugler J., Maurice D., Roegel J.C.,
RA   Briand J.P., Hoebeke J.;
RT   "Modulation of the M2 muscarinic receptor activity with monoclonal
RT   anti-M2 receptor antibody fragments.";
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ746180; CAG34081.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   255 AA;  27445 MW;  B68BD38395DF713B CRC64;

  Query Match             73.6%;  Score 936.5;  DB 2;  Length 255;
  Best Local Similarity   72.9%;  Pred. No. 1.6e-62;
  Matches  180;  Conservative   23;  Mismatches   33;  Indels   11;  Gaps    3;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              ||:||:||| |||||||||:|||||||||||||||||||||||||||||:||||||||
Db          1 QVQLQQSGGDLVKPGGSLKVSCAASGFTFSSYGMSWVRQTPDKRLEWVATITSGGSYTYY 60

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARG---NWEGWYFDVWGQGTTVTV 117
              ||||||||||||||||||||||||||||||||||||||     ::|  || ||||||:||
Db         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARHINYRYDG-AFDYWGQGTTLTV 119

Qy        118 SSGGGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSI-------RYIYWYQ 170
              |||||||||||||||||:| : |||: :| | ||:| |:| :| |:        |: |||
Db        120 SSGGGGSGGGGSGGGGSDIVMAQSPSSLSVSAGEKVIMSCKSSQSLLNSRNQKNYLAWYQ 179

Qy        171 QKPGSSPRLLIYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYT 230
              |||| ||:|||| |    ||| ||:|||||| ::|||: ::||| | ||||    || |
Db        180 QKPGQSPKLLIYGASTRESGVPDRFTGSGSGTDFTLTISSVQAEDLAVYYCQNDHSYPLT 239

Qy        231 FGGGTKL 237
              || ||||
Db        240 FGAGTKL 246


RESULT 3
Q9QYF0
ID   Q9QYF0      PRELIMINARY;      PRT;   298 AA.
AC   Q9QYF0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   CN 8 single chain antibody.
GN   Name=CN 8 scFv;
OS   synthetic construct.
OC   artificial sequences.
OX   NCBI_TaxID=32630;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20183931; PubMed=10706631;
RA   Shinohara N., Demura T., Fukuda H.;
RT   "Isolation of a vascular cell wall-specific monoclonal antibody
RT   recognizing a cell polarity by using a phage display subtraction
RT   method.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:2585-2590(2000).
DR   EMBL; AB036341; BAA88633.1; -.
DR   PIR; A33933; A33933.
DR   PIR; S19112; S19112.
DR   HSSP; P01820; 1A7O.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00406; IGv; 2.
DR   PROSITE; PS50835; IG_LIKE; 2.
SQ   SEQUENCE   298 AA;  31867 MW;  E0F96B8A17004317 CRC64;

  Query Match             65.9%;  Score 838;  DB 2;  Length 298;
  Best Local Similarity   68.5%;  Pred. No. 4.6e-55;
  Matches  163;  Conservative   24;  Mismatches   49;  Indels    2;  Gaps    2;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              |||||:|||||||||||||||||||  || | |||||| | | |||:  |:   |   |
Db         40 QVKLQQSGGGLVKPGGSLKLSCAASGSDFSRYWMSWVRQAPGKGLEWIGEINPDSSTINY 99

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
                |:| :| |||||||||||||||| ::|||||:||||| :: | :   ||||||||||||
Db        100 TPSLKDKFIISRDNAKNTLYLQMSKVRSEDTALYYCARASYYG-HSAYWGQGTTVTVSSG 158

Qy        121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSI-RYIYWYQQKPGSSPRL 179
              |||||||||||||:|||||||| :||| || ||:|| || :|  |: ||||| | ||:|
Db        159 GGGSGGGGSGGGGSDIELTQSPASLSASVGETVTITCRASGNIHNYLAWYQQKQGKSPQL 218

Qy        180 LIYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
              |:|:   :| ||| |||||||||| ||| || :: ||  :|||| :   ||||||||||
Db        219 LVYNAKTLADGVPSRFSGSGSGTQYSLKINSLQPEDFGSYYCQHFWTTPYTFGGGTKL 276


RESULT 4
Q921A6
ID   Q921A6      PRELIMINARY;      PRT;   241 AA.
AC   Q921A6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Anti-CEA 79 single chain Fv (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98170165; PubMed=9509426;
RA   Chung J.H., Choi S.J., Kim H.J., Kim I.J., Choi I.H., Lee S.D.,
RA   Yi K.S., Suh P.G., Ryu S.H., Chung H.K.;
RT   "Cloning and characterization of cDNAs encoding VH and VL of a
RT   monoclonal anti-CEA antibody (CEA 79) cross-reactive with NCA-95 and
RT   generation of a single-chain Fv molecule (scFv).";
RL   Mol. Cells 7:816-819(1997).
DR   EMBL; U88067; AAB48044.1; -.
DR   PIR; S19965; S19965.
DR   PIR; S19967; S19967.
DR   PIR; S19968; S19968.
DR   PIR; S26325; S26325.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00406; IGv; 2.
DR   PROSITE; PS50835; IG_LIKE; 2.
FT   NON_TER       1      1
FT   NON_TER     241    241
SQ   SEQUENCE   241 AA;  26086 MW;  0276887248E9C771 CRC64;

  Query Match             55.9%;  Score 710.5;  DB 2;  Length 241;
  Best Local Similarity   58.3%;  Pred. No. 1.4e-45;
  Matches  141;  Conservative   34;  Mismatches   56;  Indels   11;  Gaps    5;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              |||||:||  | ||| ::|:||| |||:||: |||:||:| | | |:|:  |::      |
Db          1 QVKLQQSGPELKKPGETVKISCKASGYTFTDYGMNWVKQAPGKGLKWMGWINTYTGEPTY 60
```

```
Qy          61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
               |  ||||  | : : :| |||:::||:|||| |:||| :    ||| ||||||||||||
Db          61 ADDFKGRFAFSLETSASTAYLQINNLKNEDTATYFCARKDLLR-YFDYWGQGTTVTVSSG 119

Qy         121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSI-RYIYWYQQKPGSSPR- 178
               ||||||||||||||:|||||||: :||| | :||:|| ||  | :|| ||| |||  ||
Db         120 GGGSGGGGSGGGGSDIELTQSPSSLSASLGGKVTITCKASQDINKYIAWYQHKPGKGPRS 179

Qy         179 ---LLIYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGT 235
                  | ||    : ||:| ||||||||  || :|: :| || |||||  :    :||||||
Db         180 AHTLHIY----IQPGIPSRFSGSGSGRDYSFSISNLEPEDIATYYCLHYDNL-HTFGGGT 234

Qy         236 KL 237
               ||
Db         235 KL 236


RESULT 5
Q7TQM2
ID   Q7TQM2      PRELIMINARY;      PRT;   243 AA.
AC   Q7TQM2;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   ScFv 6H8 protein (Fragment).
GN   Name=scFv 6H8;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Balb/C;
RX   MEDLINE=22853226; PubMed=12860977;
RA   Peter J.C., Eftekhari P., Billiald P., Wallukat G., Hoebeke J.;
RT   "scFv single chain antibody variable fragment as inverse agonist for
RT   the beta-2 adrenergic receptor.";
RL   J. Biol. Chem. 278:36740-36747(2003).
DR   EMBL; AJ574851; CAE00495.1; -.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00406; IGv; 2.
DR   PROSITE; PS50835; IG_LIKE; 2.
FT   NON_TER       1      1
SQ   SEQUENCE   243 AA;  25976 MW;  BEFFF64D2DCF4F76 CRC64;

  Query Match             54.4%;  Score 692.5;  DB 2;  Length 243;
  Best Local Similarity   56.5%;  Pred. No. 3.2e-44;
  Matches  135;  Conservative   38;  Mismatches   59;  Indels    7;  Gaps    4;

Qy           1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
               ||:||:||  ||:|| |:|||| |||:||::| | ||:|   : |||:  |  |  | |
Db           1 QVQLQQSGSELVRPGASVKLSCKASGYTFTTYWMHWVKQRHGQGLEWIGNIYPGSGITNY 60

Qy          61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
                 :  | :  :: | : :| |: :||| |||:|:|||||   |   |||| |||:|||||
Db          61 DEKFKNKGILTVDTSSSTAYMHLSSLASEDSAVYYCARG---GRGLDVWGAGTTLTVSSG 117

Qy         121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSI-RYIYWYQQKPGSSPRL 179
               ||||||||||||||:|::||| :  | | |:|||:|| ||   |   : |||||||::|||
Db         118 GGGSGGGGSGGGGSDIQMTQSSSSFSVSLGDRVTITCKASEDIYNRLAWYQQKPGNAPRL 177

Qy         180 LIYDTSNVAPGVPFRFSGSGSGTSYSLTINRMEAEDAATYYCQE-WSGYPYTFGGGTKL 237
               ||   :::  ||| |||||||| |:|:|  :: || ||||||: ||    ||||||||
Db         178 LISGATSLETGVPSRFSGSGSGKDYTLSITSLQTEDVATYYCQQYWS--TRTFGGGTKL 234


RESULT 6
Q925S1
ID   Q925S1      PRELIMINARY;      PRT;   218 AA.
AC   Q925S1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   MRP5 (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   PubMed=11819679;
RA   Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X.,
RA   Su C.;
RT   "Mechanism of exogenous nucleic acids and their precursors improving
RT   the repair of intestinal epithelium after irradiation in mice.";
RL   World J. Gastroenterol. 6:709-717(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RA   Cui D., Zeng G., Yan X., Li X., Su C.;
RT   "Cloning of mouse genes related to repairing of intestinal epithelium
RT   of the irradiated mice by treatment with the intestinal RNA of mice of
RT   the same strain.";
RL   Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).
DR   EMBL; AF240168; AAK43733.1; -.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
FT   NON_TER     218    218
SQ   SEQUENCE   218 AA;  23013 MW;  527E4FA8F7982817 CRC64;

  Query Match             46.2%;  Score 588;  DB 2;  Length 218;
  Best Local Similarity   53.0%;  Pred. No. 2e-36;
  Matches  115;  Conservative   36;  Mismatches   60;  Indels    6;  Gaps    2;

Qy           1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
               |||||:||  | ||| ::::|| |||:||:: || ||:: | | |:|:  |::      |
Db           3 QVKLQQSGPELKKPGETVRISCKASGYTFTTAGMQWVQKMPGKGLKWIGWINTHSGVPKY 62

Qy          61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
               :  ||||  | : : :| |||:|:|||:|||| |:|  | :::|  |  ||||||||||||
Db          63 AEEFKGRFAFSLETSASTAYLQISNLKNEDTATYFCMRWDYDGG-FAYWGQGTTVTVSSG 121

Qy         121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSS-----IRYIYWYQQKPGS 175
               |||||||||| |||:| |||||| :: | |:| |::| || |     | :: |:|||||
Db         122 GGGSGGGGSGSGGSDIVLTQSPASLAVSLGQRATISCRASESVDNIGISFMNWFQQKPGQ 181

Qy         176 SPRLLIYDTSNVAPGVPFRFSGSGSGTSYSLTINRME 212
               |:|||  |    |||     |||| :|| |  ||
Db         182 PPKLLIYAASKQGSGVPAGLLASGSGTDFSLNIYPME 218


RESULT 7
Q99KA4
ID   Q99KA4      PRELIMINARY;      PRT;   487 AA.
AC   Q99KA4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   LOC380791 protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Czech II;
RC   TISSUE=Mammary tumor metastatized to lung. Tumor arose spontaneously;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
```

```
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Czech II;
RC   TISSUE=Mammary tumor metastatized to lung. Tumor arose spontaneously;
RA   Strausberg R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC004786; AAH04786.1; -.
DR   HSSP; P01810; 2FBJ.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 2.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
SQ   SEQUENCE   487 AA;  52554 MW;  7DC8E96DB333077B CRC64;

  Query Match             42.7%;  Score 543.5;  DB 2;  Length 487;
  Best Local Similarity   50.8%;  Pred. No. 1.1e-32;
  Matches  127;  Conservative   16;  Mismatches   36;  Indels   71;  Gaps    8;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              :|:| |||||||||||||||||||||||||| ||||||||:|||||||||| |||||||
Db         20 EVQLVESGGGLVKPGGSLKLSCAASGFTFSSYAMSWVRQTPEKRLEWVATISDGGSYTYY 79

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCAR---GNWEGWY--FDVWGQGTTV 115
              ||:|||||||||||||| |||||| ||||||||||||    |:  | |  || ||||||:
Db         80 PDNVKGRFTISRDNAKNNLYLQMSHLKSEDTAMYYCARDMGGSPYGGYSRFDYWGQGTTI 139

Qy        116 TVSSGGGGSGGGGSGGGGSNIELTQSPAI--------MSASPGERVTMTCSASSSIRYIY 167
              ||||                | ::| |        :|: |   | : |
Db        140 TVSS----------------ESARNPTIYPLTLPRALSSDP---VIIGC---------- 169

Qy        168 WYQQKPGSSPRLLIYD-----TSNVAPG----------VPFRFSGSGSGTSYSLTINRM 211
                        ||:|     | ||  |            |   || |   |  ||:  :
Db        170 -----------LIHDYFPSGTMNVTWGKSGKDITTVNFPPALASGGGYTMSSQLTLPAV 217

Qy        212 EAEDAATYYC 221
              |  :  :  |
Db        218 ECPEGESVKC 227


RESULT 8
Q920E7
ID   Q920E7      PRELIMINARY;      PRT;   119 AA.
AC   Q920E7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Pterin-mimicking anti-idiotope heavy chain variable region
DE   (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Atkin J.D., Iape A., Jennings I.G., Horaitis O., Cotton R.G.H.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF307937; AAL09421.1; -.
DR   PIR; C25913; C25913.
DR   HSSP; P01783; 1IGC.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
FT   NON_TER       1      1
FT   NON_TER     119    119
SQ   SEQUENCE   119 AA;  13025 MW;  F6E904044381CA7C CRC64;

  Query Match             41.7%;  Score 531;  DB 2;  Length 119;
  Best Local Similarity   86.9%;  Pred. No. 2e-32;
  Matches  106;  Conservative    3;  Mismatches    7;  Indels    6;  Gaps    2;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              :|:| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EVQLVESGGDLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWY---FDVWGQGTTVTV 117
              ||||||||||||||||||||||||||||||||||||||||||||    | |   |  ||||| |||
Db         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCAR---HGDYDVGFAYWGQGTLVTV 117

Qy        118 SS 119
              |:
Db        118 SA 119


RESULT 9
Q91Z07
ID   Q91Z07      PRELIMINARY;      PRT;   486 AA.
AC   Q91Z07;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   LOC380791 protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Czech II;
RC   TISSUE=Mammary tumor metastatized to lung. MMTV-LTR/Wnt1 model.
RC   Expression driven by an MMTV-LTR enhancer.;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Czech II;
RC   TISSUE=Mammary tumor metastatized to lung. MMTV-LTR/Wnt1 model.
RC   Expression driven by an MMTV-LTR enhancer.;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC010324; AAH10324.1; -.
DR   HSSP; P01789; 1MCP.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 2.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
SQ   SEQUENCE   486 AA;  52681 MW;  4FEF835125DA870B CRC64;

  Query Match             39.5%;  Score 502;  DB 2;  Length 486;
  Best Local Similarity   46.8%;  Pred. No. 1.4e-29;
  Matches  117;  Conservative   21;  Mismatches   40;  Indels   72;  Gaps    8;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              :| | |||||||||||||||||  |||:|:|| ||||||||::|||||| |:|||: |||
Db         20 EVHLVESGGGLVKPGGSLKLSCVVSGFSFTSYDMSWVRQTPERRLEWVAAITSGGN-TYY 78

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGN-----WEGWYFDVWGQGTTV 115
              ||:||||||:|||||| |||||||||||||||||| |       : | ||| ||||||:
Db         79 PDNVKGRFTVSRDNAKYTLYLQMSSLKSEDTAMYYCVRPEIPIYYYSGSYFDSWGQGTTI 138

Qy        116 TVSSGGGGSGGGGSGGGGSNIELTQSPAI--------MSASPGERVTMTCSASSSIRYIY 167
              ||||                  |  ::| |        :|: |   | : |
Db        139 TVSS-----------------ESARNPTIYPLTLPRALSSDP---VIIGC---------- 168

Qy        168 WYQQKPGSSPRLLIYD-----TSNVAPG-----------VPFRFSGSGSGTSYSLTINRM 211
                          ||:|     | ||  |           |   || |   |  ||:  :
Db        169 -----------LIHDYFPSGTMNVTWGKSGKDITTVNFPPALASGGGYTMSSQLTLPAV 216

Qy        212 EAEDAATYYC 221
              |  :  :  |
Db        217 ECPEGESVKC 226


RESULT 10
Q91XE1
ID   Q91XE1      PRELIMINARY;      PRT;   480 AA.
AC   Q91XE1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Igh-VJ558 protein (Fragment).
GN   Name=Igh-VJ558;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Colon;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Colon;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC010798; AAH10798.1; -.
DR   HSSP; P01789; 1MCP.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 2.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
FT   NON_TER      1      1
SQ   SEQUENCE   480 AA;  51936 MW;  20B9234EEF2B41ED CRC64;

  Query Match             39.0%;  Score 496;  DB 2;  Length 480;
  Best Local Similarity   48.1%;  Pred. No. 3.9e-29;
  Matches  115;  Conservative   20;  Mismatches   46;  Indels   58;  Gaps    6;

Qy          2 VKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYYP 61
              ||| |||||||||||||:|||||||| ||:  ||||||||:||||||||: |  |:||
Db         20 VKLVESGGGLVKPGGSLRLSCAASGFIFSNSYMSWVRQTPEKRLEWVATISNSGYATHYP 79

Qy         62 DSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSGG 121
              ||:||||||||||:||: |||:|| |||||:||| ||::  ||||||| ||||||||
Db         80 DSMKGRFTISRDNAQNTVLLQMTSLNSEDTAVYYCTRGDY--WYFDVWGAGTTVTVSS-- 135

Qy        122 GGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRL-- 179
                                 | : | |                  |   : |   | | :
Db        136 ---------------EPAREPTI-------------------YPLTFPQALSSDPVIIG 160

Qy        180 -LIYD-----TSNVAPG-----------VPFRFSGSGSGTSYSLTINRMEAEDAATYYC 221
               ||:|     | ||  |           |   ||     |  ||:  :|  :  :  |
Db        161 CLIHDYFPSGTMNVTWGKSGKDITTVNFPPALASGGRYTMSSQLTLPAVECPEGESVKC 219


RESULT 11
AAH24289
ID   AAH24289    PRELIMINARY;      PRT;   471 AA.
AC   AAH24289;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
```

```
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RA   Strausberg R.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC024289; AAH24289.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   471 AA;  51791 MW;  388F7F4CF588660E CRC64;

  Query Match             38.3%;  Score 487;  DB 2;  Length 471;
  Best Local Similarity   46.3%;  Pred. No. 1.8e-28;
  Matches  105;  Conservative   33;  Mismatches   61;  Indels   28;  Gaps    4;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              :|:| |||||||||||||:|||||||||||||| |:|||| | | ||||:::||  || ||
Db         20 EVQLVESGGGLVKPGGSLRLSCAASGFTFSSYSMNWVRQAPGKGLEWVSSMSSSSSYIYY 79

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWE---GWYFDVWGQGTTVTV 117
               ||||||||||||||:|||||:||::||||:|||||    :    ||||:||:|| |||
Db         80 ADSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARDLRQLTSYWYFDLWGRGTLVTV 139

Qy        118 SSGGGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSP 177
              ||                    |: |::   :|  : |   :|:       :: :     |
Db        140 SSAS----------------TKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPE-----P 177

Qy        178 RLLIYDTSNVAPGV---PFRFSGSGSGTSYSLTINRMEAEDAATYYC 221
                : :::  :  ||   |      ||  :  |:      :    || |
Db        178 VTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC 224


RESULT 12
Q91WP5
ID   Q91WP5      PRELIMINARY;      PRT;   479 AA.
AC   Q91WP5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Igh-VJ558 protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Colon;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Colon;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC013656; AAH13656.1; -.
DR   HSSP; P01789; 1MCP.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 2.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
SQ   SEQUENCE   479 AA;  51603 MW;  ECB2D0877748584F CRC64;

  Query Match             38.3%;  Score 487;  DB 2;  Length 479;
  Best Local Similarity   47.1%;  Pred. No. 1.8e-28;
  Matches  113;  Conservative   19;  Mismatches   48;  Indels   60;  Gaps    6;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              :|:| |||||||||||||:|||||| |||:| |||||:|:|||||| |:| |  |||
Db         20 EVQLVESGGGLVKPGGSLKVSCAASGLTFSNYAMSWVRQSPEKRLEWVAAINSNGGNTYY 79

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWYFDVWGQGTTVTVSSG 120
               |::||||||||||:||||||||:||||| ||| ||     ||||| || |||||
Db         80 SDTMKGRFTISRDNAKSTLYLQMSSLRSEDTAFYYCVRGG----YFDVWGAGTAVTVSS- 134

Qy        121 GGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRL- 179
                           |  : | |                 |   :|   | | :
Db        135 ---------------EPAREPTI-------------------YPLTFPQALSSDPVII 158

Qy        180 --LIYD-----TSNVAPG----------VPFRFSGSGSGTSYSLTINRMEAEDAATYYC 221
                ||:|     | ||  |           |   ||     |  ||:  :|  :  :  |
Db        159 GCLIHDYFPSGTMNVTWGKSGKDITTVNFPPALASGGRYTMSSQLTLPAVECPEGESVKC 218


RESULT 13
AAR11048
ID   AAR11048    PRELIMINARY;      PRT;   108 AA.
AC   AAR11048;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   ANA immunoglobulin kappa light chain (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B6.Sle1; TISSUE=Spleen;
RA   Liang Z., Xie C., Chen C., Kreska D., Hsu K., Zhou J.X., Mohan C.;
RT   "Antinuclear autoantibodies from B6.Sle1 mice.";
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY436888; AAR11048.1; -.
FT   NON_TER       1      1
FT   NON_TER     108    108
SQ   SEQUENCE   108 AA;  11738 MW;  C11222F2FDFBC160 CRC64;

  Query Match             37.6%;  Score 478;  DB 2;  Length 108;
  Best Local Similarity   87.0%;  Pred. No. 1.7e-28;
  Matches   87;  Conservative    9;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        138 LTQSPAIMSASPGERVTMTCSASSSIRYIYWYQQKPGSSPRLLIYDTSNVAPGVPFRFSG 197
              :||:||||||||||:||||||||||: |:|||||||||||||||||||:| ||| ||||
```

```
Db          1 MTQTPAIMSASPGEKVTMTCSASSSVSYMYWYQQKPGSSPRLLIYDTSNLASGVPVRFSG 60

Qy        198 SGSGTSYSLTINRMEAEDAATYYCQEWSGYPYTFGGGTKL 237
              |||||||||||:|||||||||||||:|| ||:||||||||
Db         61 SGSGTSYSLTISRMEAEDAATYYCQQWSSYPWTFGGGTKL 100


RESULT 14
Q6GMY2
ID   Q6GMY2      PRELIMINARY;      PRT;   606 AA.
AC   Q6GMY2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Primary B-Cells;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Primary B-Cells;
RA   Strausberg R.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC073758; AAH73758.1; -.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 4.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 4.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 5.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW   Hypothetical protein.
SQ   SEQUENCE   606 AA;  66184 MW;  B6B38B51114E4C55 CRC64;

  Query Match             37.5%;  Score 477.5;  DB 2;  Length 606;
  Best Local Similarity   68.3%;  Pred. No. 1.2e-27;
  Matches   95;  Conservative   12;  Mismatches   17;  Indels   15;  Gaps    2;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISSGGSYTYY 60
              ||:| |||||||||||||:|||||||||| | |||:|| | | ||||: ||| |||| |
Db         20 QVQLVESGGGLVKPGGSLRLSCAASGFTFSDYYMSWIRQAPGKGLEWVSYISSSSSYTNY 79

Qy         61 PDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGN----------WEGWYF--- 106
                  ||||||||||||||:|||||:||::||||:||||||           | :|:
Db         80 ADSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARGGNGIAAAGRVVYAEDYYYYG 139

Qy        107 -DVWGQGTTVTVSSGGGGS 124
               ||||||||||||| :
Db        140 MDVWGQGTTVTVSSGSASA 158


RESULT 15
Q6MZU6
ID   Q6MZU6      PRELIMINARY;      PRT;   464 AA.
AC   Q6MZU6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein DKFZp686C15213.
GN   Name=DKFZp686C15213;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Human rectum tumor;
RG   THE GERMAN HUMAN CDNA CONSORTIUM;
RA   Bloecker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA   Fobo G., Han M., Wiemann S.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX640874; CAE45931.1; -.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   464 AA;  51099 MW;  2FCA72C66E8A0ABC CRC64;

  Query Match             37.3%;  Score 475;  DB 2;  Length 464;
  Best Local Similarity   59.8%;  Pred. No. 1.4e-27;
  Matches   98;  Conservative   18;  Mismatches   20;  Indels   28;  Gaps    5;

Qy          1 QVKLQESGGGLVKPGGSLKLSCAASGFTFSSYGMSWVRQTPDKRLEWVATISS-GGSYTY 59
              :|:| |||||||||||||:|||||||||||| |:|||| | | ||||:: || |||| |
Db         20 EVQLVESGGGLVKPGGSLRLSCAASGFTFSSYSMNWVRQAPGKGLEWVSSFSSRGGSYEY 79

Qy         60 YPDSVKGRFTISRDNAKNTLYLQMSSLKSEDTAMYYCARGNWEGWY-FDVWGQGTTVTVS 118
              | |||||||||||||||:|||||:||::|:||:|||||    ||:  | ||||| ||||
Db         80 YADSVKGRFTISRDNAKNSLYLQMNSLRAENTAVYYCARD--LGWFGLDYWGQGTLVTVS 137

Qy        119 SGGGGSGGGGSGGGGSNIELTQSPAIMSASPGERVTMTCSASSS 162
              |                    |: |::   :|       || |:|
Db        138 SAS----------------TKGPSVFPLAP-------CSRSTS 157


Search completed: December  8, 2004, 17:09:58
Job time : 189.085 secs
```